Walmart Customer Trip Type Classification

1. 분석목적 (Purpose of Analysis)

Walmart Recruiting: Trip Type Classification 이용, Walmart 소비자 패턴을 분석하여 범주화된 Trip Type에 해당되는 VisitNumber 예측 및 분류.

2. 사용한 라이브러리

```
## Loading required package: grid
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: modeltools
```

```
## Loading required package: stats4
```

```
## Loading required package: strucchange
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: sandwich
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```
## Loading required package: Hmisc
```

```
## Loading required package: survival
```

```
##
## Attaching package: 'survival'
```

```
## The following object is masked from 'package:caret':
##
##     cluster
```

```
## Loading required package: Formula
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following object is masked from 'package:e1071':
##
##     impute
```

```
## The following objects are masked from 'package:dplyr':
##
##     combine, src, summarize
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units
```

```
## funModeling v.1.6.5 :)
## Examples and tutorials at livebook.datascienceheroes.com
```

```
##
## Attaching package: 'kknn'
```

```
## The following object is masked from 'package:caret':
##
##     contr.dummy
```

```
## Type 'citation("pROC")' for a citation.
```

```
##
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':
##
##     cov, smooth, var
```

```
## randomForest 4.6-12
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:Hmisc':
##
##     combine
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

수정 전 데이터

```
wal_dataset <- read.csv("C:/Users/rudghksldl/Desktop/train 3.csv")

head(wal_dataset, 20)
```

```
##    TripType VisitNumber Weekday         Upc ScanCount
## 1       999           5  Friday 68113152929        -1
## 2        30           7  Friday 60538815980         1
## 3        30           7  Friday  7410811099         1
## 4        26           8  Friday  2238403510         2
## 5        26           8  Friday  2006613744         2
## 6        26           8  Friday  2006618783         2
## 7        26           8  Friday  2006613743         1
## 8        26           8  Friday  7004802737         1
## 9        26           8  Friday  2238495318         1
## 10       26           8  Friday  2238400200        -1
## 11       26           8  Friday  5200010239         1
## 12       26           8  Friday 88679300501         2
## 13       26           8  Friday 22006000000         1
## 14       26           8  Friday  2236760452         1
## 15       26           8  Friday 88679300501        -1
## 16       26           8  Friday  2238400200         2
## 17       26           8  Friday  3019294203         1
## 18       26           8  Friday 72450408840         1
## 19       26           8  Friday 25541500000         2
## 20       26           8  Friday  2310010776         1
##    DepartmentDescription FinelineNumber
## 1     FINANCIAL SERVICES           1000
## 2                  SHOES           8931
## 3          PERSONAL CARE           4504
## 4   PAINT AND ACCESSORIES          3565
## 5   PAINT AND ACCESSORIES          1017
## 6   PAINT AND ACCESSORIES          1017
## 7   PAINT AND ACCESSORIES          1017
## 8   PAINT AND ACCESSORIES          2802
## 9   PAINT AND ACCESSORIES          4501
## 10  PAINT AND ACCESSORIES          3565
## 11            DSD GROCERY          4606
## 12  PAINT AND ACCESSORIES          3504
## 13 MEAT - FRESH & FROZEN           6009
## 14  PAINT AND ACCESSORIES             7
## 15  PAINT AND ACCESSORIES          3504
## 16  PAINT AND ACCESSORIES          3565
## 17  PAINT AND ACCESSORIES          2801
## 18  PAINT AND ACCESSORIES          1028
## 19                  DAIRY          1305
## 20      PETS AND SUPPLIES          3300
```

변수 설명

TripType - a categorical id representing the type of shopping trip the customer made. This is the ground truth that you are predicting. TripType_999 is an “other” category. VisitNumber - an id corresponding to a single trip by a single customer
Weekday - the weekday of the trip
Upc - the UPC number of the product purchased
ScanCount - the number of the given item that was purchased. A negative value indicates a product return.
DepartmentDescription - a high-level description of the item’s department
FinelineNumber - a more refined category for each of the products, created by Walmart

변수 변환

```
dt <- na.omit(read.csv("C:/Users/rudghksldl/Desktop/newdata (1).csv"))
dt <- dt[ ,c(-1,-2,-5,-6)]
dt$triptype <- as.factor(dt$triptype)
glimpse(dt)
```

```
## Observations: 94,247
## Variables: 74
## $ num_prod                     <int> -1, 2, 28, 3, 3, 4, 7, 9, 4, 9, 3,...
## $ weekday                      <fctr> Friday, Friday, Friday, Friday, F...
## $ TotalDescCount               <int> 1, 2, 22, 3, 3, 4, 7, 8, 4, 9, 2, ...
## $ max_prodrate                 <dbl> 1.000, 0.500, 0.727, 0.667, 0.667,...
## $ X1.HR.PHOTO                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ ACCESSORIES                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0...
## $ AUTOMOTIVE                   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ BAKERY                       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ BATH.AND.SHOWER              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ BEAUTY                       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ BEDDING                      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ BOOKS.AND.MAGAZINES          <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ BOYS.WEAR                    <int> 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0...
## $ BRAS...SHAPEWEAR             <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ CAMERAS.AND.SUPPLIES         <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ CANDY..TOBACCO..COOKIES      <int> 0, 0, 0, 0, 1, 0, 0, 0, 2, 0, 0, 0...
## $ CELEBRATION                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ COMM.BREAD                   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ CONCEPT.STORES               <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ COOK.AND.DINE                <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ DAIRY                        <int> 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ DSD.GROCERY                  <int> 0, 0, 1, 0, 2, 1, 0, 0, 2, 0, 0, 0...
## $ ELECTRONICS                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ FABRICS.AND.CRAFTS           <int> 0, 0, 0, 0, 0, 0, 0, 8, 0, 1, 0, 0...
## $ FINANCIAL.SERVICES           <int> 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ FROZEN.FOODS                 <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0...
## $ FURNITURE                    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ GIRLS.WEAR..4.6X..AND.7.14   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ GROCERY.DRY.GOODS            <int> 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0...
## $ HARDWARE                     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ HEALTH.AND.BEAUTY.AIDS       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ HOME.DECOR                   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ HOME.MANAGEMENT              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1...
## $ HORTICULTURE.AND.ACCESS      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ HOUSEHOLD.CHEMICALS.SUPP     <int> 0, 0, 1, 0, 0, 0, 2, 0, 0, 0, 0, 0...
## $ HOUSEHOLD.PAPER.GOODS        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ IMPULSE.MERCHANDISE          <int> 0, 0, 0, 1, 0, 2, 0, 0, 0, 2, 0, 0...
## $ INFANT.APPAREL               <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ INFANT.CONSUMABLE.HARDLINES <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ JEWELRY.AND.SUNGLASSES       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0...
## $ LADIES.SOCKS                 <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ LADIESWEAR                   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ LARGE.HOUSEHOLD.GOODS        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ LAWN.AND.GARDEN              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ LIQUOR.WINE.BEER             <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ MEAT...FRESH...FROZEN        <int> 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ MEDIA.AND.GAMING             <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ MENS.WEAR                    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0...
## $ MENSWEAR                     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ NULL.                        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ OFFICE.SUPPLIES              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ OPTICAL...FRAMES             <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
```

```
## $ OPTICAL...FRAMES          <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ OPTICAL...LENSES          <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ OTHER.DEPARTMENTS         <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PAINT.AND.ACCESSORIES     <int> 0, 0, 16, 0, 0, 0, 0, 0, 0, 0, 0, ...
## $ PERSONAL.CARE             <int> 0, 1, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0...
## $ PETS.AND.SUPPLIES         <int> 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PHARMACY.OTC              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PHARMACY.RX               <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PLAYERS.AND.ELECTRONICS   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PLUS.AND.MATERNITY        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PRE.PACKED.DELI           <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ PRODUCE                   <int> 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 1...
## $ SEAFOOD                   <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ SEASONAL                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ SERVICE.DELI              <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0...
## $ SHEER.HOSIERY             <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ SHOES                     <int> 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0...
## $ SLEEPWEAR.FOUNDATIONS     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ SPORTING.GOODS            <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ SWIMWEAR.OUTERWEAR        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ TOYS                      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ WIRELESS                  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0...
## $ triptype                  <fctr> 999, 30, 26, 8, 8, 35, 41, 21, 6,...
```

새로 만든 변수 설명

id - 고객 고유값
num_prod - 구매 + 환불(-)
weekday - 방문요일
RefundCount - 고객의 총 환불량
TotalDescCount - 고객의 총 구매량
MaxDescCount - 최대값을 갖는 DepartmentDescription의 양
max_prodrate - 최대값을 갖는 DepartmentDescription의 해당비율

DepartmentDescription의 dummy 변수(one hot 인코딩) - 68열

TripType - 예측해야할 타겟 변수

4. Model Accuracy 판단

```
# train / validation data
trn_idx <- sample(1:dim(dt)[1], round(0.7*dim(dt)[1]))
w_trn <- dt[trn_idx, ]
w_val <- dt[-trn_idx, ]
```

4.1 KNN

찾은 K 값으로 모델 적합

```
w_knn <- kknn(triptype ~ .,
              w_trn,
              w_val,
              k = 7,
              distance = 2,
              kernel = "rectangular")
```

모델 적합 후 정확도 확인

```
pima_knn_fit <- fitted(w_knn)
pima_knn_cf <- confusionMatrix(w_val$triptype, as.factor(pima_knn_fit))
pima_knn_cf
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    3    4    5    6    7    8    9   12   14   15   18   19
##          3 1020    0    0    1    0   29    7    0    0    0    0    0
##          4    1    2   78    0    1    7    2    0    0    0    0    0
##          5    6    5  731    6   22   71   40    0    0    1    2    1
##          6    1    0    1  274    6   58    3    0    0    0    0    1
##          7    3    0   10    8 1208  220   24    1    0    3    2    0
```

```
##         8     7     0   111    70   194  2917   133     0     0     0     0     0
##         9    20     0    15     2    12   140  2102     3     0     9    16    31
##         12    0     0     0     1     9     3    10     3     0     1     2     0
##         14    0     0     0     0     0     0     0     0     0     0     0     0
##         15    2     0     3     1     8    47    49     1     0    97     1     0
##         18    0     0     0     0     0     9    49     0     0     1    66     0
##         19    0     0     0     0     0     0    43     0     0     0     1    28
##         20    1     0     2     1     1     1    69     0     0     1     0     0
##         21    0     0     2     1     1     3    46     0     0     5     2     1
##         22    2     0     2     0     0     0    91     0     0     0     1    22
##         23    2     0     0     0     0     0    22     0     0     0     0     0
##         24    5     0     3     4     7     9   151     0     0     3     1     1
##         25    8     0     3     6     5     7   175     2     0     0     2     0
##         26    2     0     1     0     1     1    57     0     0     1     0     0
##         27    3     0     1     2     2     2    55     3     0     1     2     0
##         28    1     0     2     0     2     0    63     0     0     0     0     0
##         29    3     0     4     0     0     2    51     0     0     0    13     1
##         30    2     0     2     1     2     2   123     0     0     0     0     0
##         31    4     0     0     1     0     4    36     0     0     0     0     0
##         32    5     0     6     3     9    44    31     1     0     3     6     0
##         33    3     0     0     1     6    62    11     0     0     0     0     0
##         34    1     0     2     4     4    37    12     0     0     0     0     0
##         35   11     0    12     7    34   140    28     0     0     2     4     1
##         36   10     0    44     4     7   115    35     0     0     3     2     0
##         37    6     0     8     1   126    71    13     0     0     5     1     4
##         38    2     0     5     7    80   107    20     0     0     0     2     0
##         39   21     0   115    72   292   191   130     8     0    22    13     4
##         40    0     1    14     3    22     2     4     3     0     4     4     0
##         41    1     0     1     0     0     0    36     0     0     1     3     3
##         42    2     0     1     3     2     7    85     1     0    26    18     3
##         43    7     0     5     1     2    11    62     3     0     8     6     0
##         44    0     0     8     4     3     2    12     0     0    10     6     1
##         999 221     2    39     7    54   263   374     0     0    12     6     7
##           Reference
## Prediction   20    21    22    23    24    25    26    27    28    29    30    31
##         3     1     0     5     1     6     0     0     1     0     1     2     0
##         4     0     0     0     0     0     0     0     0     1     1     0     0
##         5     1     3     1     0    10     6     1     2     1     1     4     1
##         6     1     1     1     0     0     0     0     0     0     0     0     0
##         7     1     1     1     0     5     5     0     3     1     1     4     0
##         8     2     1     0     0     7     3     3     2     0     1     5     2
##         9    13     5    41    16    72    60    11    15    21     7    32    54
##         12    0     1     0     0     2     2     0    14     0     2     0     0
##         14    0     1     1     0     0     0     0     0     0     0     0     0
##         15    0     7     4     0     4     5     1     1     2     0     1     1
##         18    4     4     0     0     4     2     0     0     0    25     1     0
##         19    0     0    44     1     0     0     0     0     1     0     0     0
##         20  123     0     1     0     0     1     0     0     0     0     0     0
##         21    0   111     0     0     2     0     1     1     0     1     0     0
##         22    0     1    99     7     4     4     1     0     0     0     1     8
##         23    0     0     2    17     0     0     0     0     0     0     0     0
##         24    2     6     2     0   463    10     5     3     2     0     2     0
##         25    3     2     3     1    12   775     4     3     1     0    21     1
##         26    2     2     0     0     5     0    45     0     1     1     0     0
##         27    0     1     0     0     4     0     2   144     0     0     0     0
##         28    0     1     0     0     6     2     0     1    65    16     1     0
##         29    1     0     0     0     1     4     0     0    16     8     1     0
##         30    1     2     1     0     1    27     0     1     1     0   126     1
##         31    0     0     1     1     0     0     0     0     1     0     0   139
##         32    0     5     0     0     3    10     0     1     1     1     9     0
##         33    3     3     1     0     7     3     2     2     0     2     0     1
##         34    1     0     1     0     1     0     1     1     0     0     0     0
##         35    2     3     0     0     5     3     1     1     2     1     1     0
##         36    3     7     5     0    14    22     2    11     1     0     5     1
##         37    4     0     5     0    10     7     2     1     3     0     4     0
##         38    2     1     1     0     3     5     1     3     0     1     6     0
##         39   27     9    30     2    68    72    20    19    10     9    34     3
##         40    2     5     2     1    21    46     7     7     4     0     2     2
##         41    2     0     9     0    16    37     7     1     1     2    28     1
##         42   15    29    29     3    49    96    17    27     9     4    24     4
##         43   10    21     1     0    21    22     9     8     3     1    10     1
##         44    2    13     9     1    33    71    10    17     3     2    14     2
##         999   9     2    23     4    59    66     4    11    13     3    11    10
```

```
##        999     9     2    23     4    59    66     4    11    15     3    11    10
##           Reference
## Prediction   32   33   34   35   36   37   38   39   40   41   42   43
##        3      0    0    0    0    0    0    0    1    0    0    0    0
##        4      0    1    0    0    2    0    0    2    0    0    0    0
##        5      0    1    2    5   22    2    4   28    0    0    1    1
##        6      0    0    0    4    1    0    1   13    1    0    0    0
##        7      2    2    1    3    1   47   36   86    1    0    0    0
##        8     43   15   11   24   38   14   24   16    0    0    0    0
##        9     25    1    0    3    4    0    1   10    0    0    3    1
##        12     0   10    1    2    2    4    1   10    5    0    0    0
##        14     0    0    0    0    0    0    0    0    0    0    0    0
##        15     2   10    0    7    3    2    2   29    3    0    1    0
##        18     3    0    0    3    3    1    0    2    0    0    2    0
##        19     0    0    1    0    0    0    0    0    0    1    0    0
##        20     0    0    1    0    1    0    0    0    0    1    2    1
##        21     0    1    0    2    3    0    0    0    0    0    6    3
##        22     0    2    2    0    0    0    0    5    3    1    0    0
##        23     0    0    0    0    0    0    0    2    0    0    0    0
##        24     1    9    2    4    5    3    1   25    8    1    8    4
##        25     5    2    1    5    7    3    1   17    2    1    8    2
##        26     0    0    0    1    0    0    2    5    1    1    0    0
##        27     0    0    4    0    1    1    0    3    0    0    0    0
##        28     1    1    0    0    1    0    0    1    0    0    0    1
##        29     2    0    1    0    0    2    0    8    0    0    4    0
##        30     0    1    1    1    4    1    0    7    0    0    3    0
##        31     0    0    0    0    0    0    0    1    0    0    0    0
##        32   382    7    2    5    5    2    8   22    5    0    3    0
##        33     2  207    0    6   11    4    1   54    0    0    0    1
##        34     0    0  126    1    2    0    1   19    6    1    0    0
##        35     4    5    1  282    8   13   17   49    3    0    1    0
##        36     5    4    1    7  459    3    4   78    7    0    2    1
##        37     1    1    5    7    2  398   12   54   47    0    1    1
##        38     8    7    3   37    6   30  336  146   47    0    1    1
##        39    51   82   46  110  142   56  158 1110   26    1    6    5
##        40    11   27   21   16   34   98   88  456  914    0    2    5
##        41     6    1    0    2    1    0    0    3    0    2    6    2
##        42    24    1    2    6    3    0    1   11    3    4   38    2
##        43     4    0    1    1    9    0    1   16    0    1    3    4
##        44    13    3    5    0   10    1    1   39   11    3   17    4
##        999   20    3    4    3   17    6   15    9    0    1    3    0
##           Reference
## Prediction   44  999
##        3      0   23
##        4      0    0
##        5      0    5
##        6      1    1
##        7      0    4
##        8      0   20
##        9      0   92
##        12     0    2
##        14     0    0
##        15     0    2
##        18     0    1
##        19     0    8
##        20     0    1
##        21     0    0
##        22     0   18
##        23     0    2
##        24     0   37
##        25     4   28
##        26     0    7
##        27     0    2
##        28     0    2
##        29     0    1
##        30     0    6
##        31     0    1
##        32     0    1
##        33     0    2
##        34     0    1
##        35     1    1
##        36     0    7
##        37     0    1
```

```
##        38     0   10
##        39     0    9
##        40     9    0
##        41     3    2
##        42     7    8
##        43     0    1
##        44     7    2
##        999    0 1240
##
## Overall Statistics
##
##                Accuracy : 0.5683
##                  95% CI : (0.5625, 0.5741)
##     No Information Rate : 0.1621
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.5371
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: 3  Class: 4 Class: 5 Class: 6 Class: 7
## Sensitivity           0.73753 2.000e-01  0.59383 0.552419  0.56927
## Specificity           0.99710 9.966e-01  0.99053 0.996580  0.98180
## Pos Pred Value        0.92896 2.041e-02  0.74063 0.742547  0.71734
## Neg Pred Value        0.98664 9.997e-01  0.98168 0.992044  0.96563
## Prevalence            0.04891 3.537e-04  0.04354 0.017543  0.07505
## Detection Rate        0.03608 7.074e-05  0.02585 0.009691  0.04272
## Detection Prevalence  0.03883 3.466e-03  0.03491 0.013051  0.05956
## Balanced Accuracy     0.86731 5.983e-01  0.79218 0.774500  0.77554
##                      Class: 8 Class: 9 Class: 12 Class: 14 Class: 15
## Sensitivity            0.6363  0.49412 0.1034483        NA  0.442922
## Specificity            0.9685  0.96940 0.9970260 9.999e-01  0.992907
## Pos Pred Value         0.7963  0.74092 0.0344828        NA  0.327703
## Neg Pred Value         0.9323  0.91540 0.9990776        NA  0.995639
## Prevalence             0.1621  0.15046 0.0010257 0.000e+00  0.007746
## Detection Rate         0.1032  0.07434 0.0001061 0.000e+00  0.003431
## Detection Prevalence   0.1296  0.10034 0.0030770 7.074e-05  0.010469
## Balanced Accuracy      0.8024  0.73176 0.5502371        NA  0.717915
##                      Class: 18 Class: 19 Class: 20 Class: 21 Class: 22
## Sensitivity           0.362637 0.2568807  0.518987  0.447581  0.306502
## Specificity           0.995942 0.9964495  0.996968  0.997110  0.993739
## Pos Pred Value        0.366667 0.2187500  0.591346  0.578125  0.361314
## Neg Pred Value        0.995871 0.9971221  0.995938  0.995121  0.992000
## Prevalence            0.006437 0.0038551  0.008382  0.008771  0.011424
## Detection Rate        0.002334 0.0009903  0.004350  0.003926  0.003501
## Detection Prevalence  0.006366 0.0045271  0.007357  0.006791  0.009691
## Balanced Accuracy     0.679290 0.6266651  0.757978  0.722345  0.650120
##                      Class: 23 Class: 24 Class: 25 Class: 26 Class: 27
## Sensitivity          0.3090909   0.50436   0.56735  0.286624  0.478405
## Specificity          0.9989369   0.98816   0.98718  0.996764  0.996818
## Pos Pred Value       0.3617021   0.58831   0.69196  0.330882  0.618026
## Neg Pred Value       0.9986538   0.98345   0.97824  0.996020  0.994401
## Prevalence           0.0019453   0.03247   0.04831  0.005553  0.010646
## Detection Rate       0.0006013   0.01638   0.02741  0.001592  0.005093
## Detection Prevalence 0.0016623   0.02783   0.03961  0.004810  0.008241
## Balanced Accuracy    0.6540139   0.74626   0.77726  0.641694  0.737612
##                      Class: 28 Class: 29 Class: 30 Class: 31 Class: 32
## Sensitivity           0.396341 0.0879121  0.361032  0.599138   0.62114
## Specificity           0.996371 0.9959195  0.993160  0.998217   0.99284
## Pos Pred Value        0.389222 0.0650407  0.397476  0.735450   0.65862
## Neg Pred Value        0.996478 0.9970516  0.992023  0.996689   0.99159
## Prevalence            0.005800 0.0032185  0.012343  0.008205   0.02175
## Detection Rate        0.002299 0.0002829  0.004456  0.004916   0.01351
## Detection Prevalence  0.005906 0.0043503  0.011212  0.006685   0.02051
## Balanced Accuracy     0.696356 0.5419158  0.677096  0.798677   0.80699
##                      Class: 33 Class: 34 Class: 35 Class: 36 Class: 37
## Sensitivity           0.512376  0.514286  0.515539   0.56877   0.57598
## Specificity           0.993254  0.996575  0.986980   0.98507   0.98539
## Pos Pred Value        0.524051  0.567568  0.438569   0.52819   0.49688
## Neg Pred Value        0.992934  0.995758  0.990409   0.98730   0.98933
## Prevalence            0.014289  0.008665  0.019346   0.02854   0.02444
## Detection Rate        0.007321  0.004456  0.009974   0.01623   0.01408
```

```
## Detection Prevalence  0.013970  0.007852  0.022742   0.03073   0.02833
## Balanced Accuracy     0.752815  0.755430  0.751260   0.77692   0.78068
##                     Class: 38 Class: 39 Class: 40 Class: 41 Class: 42
## Sensitivity            0.46927   0.47497   0.83623 1.053e-01  0.314050
## Specificity            0.98033   0.92817   0.96604 9.938e-01  0.981316
## Pos Pred Value         0.38269   0.37336   0.49755 1.130e-02  0.067376
## Neg Pred Value         0.98613   0.95150   0.99323 9.994e-01  0.997005
## Prevalence             0.02532   0.08266   0.03866 6.720e-04  0.004280
## Detection Rate         0.01188   0.03926   0.03233 7.074e-05  0.001344
## Detection Prevalence   0.03105   0.10515   0.06497 6.260e-03  0.019948
## Balanced Accuracy      0.72480   0.70157   0.90114 5.495e-01  0.647683
##                     Class: 43 Class: 44 Class: 999
## Sensitivity          0.1025641 0.2187500    0.80103
## Specificity          0.9911812 0.9882445    0.95207
## Pos Pred Value       0.0158103 0.0206490    0.49187
## Neg Pred Value       0.9987509 0.9991051    0.98804
## Prevalence           0.0013794 0.0011318    0.05475
## Detection Rate       0.0001415 0.0002476    0.04386
## Detection Prevalence 0.0089482 0.0119898    0.08916
## Balanced Accuracy    0.5468726 0.6034972    0.87655
```

4.2 RandomForest

최적의 파라미터를 찾기 위한 그리드 서치

```r
## 하이퍼 파라미터 정의
ntree = c(10, 20)
mtry = c(3, 5, 10)


# 결과값 넣을 메트릭스
tree_result <- matrix(0, length(ntree)*length(mtry),6)
iter_cnt = 1
i = 1
j = 1


# 위 파라미터들 다 넣는 포문
for(i in 1:length(ntree)){
  for(j in 1:length(mtry)){
    cat("ntree : ", ntree[i],
        ", mtry : ", mtry[j],
        "\n")


    ## 위 파라미터로 RF 모형 적합

    tmp_rf <- randomForest(triptype ~ .,
                           data = w_trn,
                           ntree = ntree[i],
                           mtry = mtry[j]
    )

    ## 위 적합으로 검증데이터 예측 수행
    tmp_tree_val_pred <- predict(tmp_rf, newdata = w_val, type = "class")


    ## 혼동행렬 작성
    tmp_tree_val_cf <- confusionMatrix(as.factor(tmp_tree_val_pred), w_val$triptype)

    ## AUROC


    ### 조건들 저장
    tree_result[iter_cnt, 1] = ntree[i]
    tree_result[iter_cnt, 2] = mtry[j]
    tree_result[iter_cnt, 3] = tmp_tree_val_cf$byClass['Recall']
    tree_result[iter_cnt, 4] = tmp_tree_val_cf$byClass['Precision']
    tree_result[iter_cnt, 5] = tmp_tree_val_cf$overall['Accuracy']
    tree_result[iter_cnt, 6] = tmp_tree_val_cf$byClass['F1']




    iter_cnt = iter_cnt +1
  }
}
```

```
## ntree :  10 , mtry :  3
## ntree :  10 , mtry :  5
## ntree :  10 , mtry :  10
## ntree :  20 , mtry :  3
## ntree :  20 , mtry :  5
## ntree :  20 , mtry :  10
```

```r
colnames(tree_result) <-  c("ntree", "mtry", "Recall", "Precision", "Accuracy",
                            "F1")


# 어큐러시 와  F1 지표를 기준으로 정렬
tree_result.df <- data.frame(tree_result)
tree_result.df <- tree_result.df[order(tree_result.df[,5],tree_result.df[,6],decreasing = T),]
head(tree_result.df)
```

```
##   ntree mtry Recall Precision  Accuracy F1
## 6    20   10     NA        NA 0.6687770 NA
## 5    20    5     NA        NA 0.6626229 NA
## 3    10   10     NA        NA 0.6569286 NA
## 2    10    5     NA        NA 0.6523661 NA
## 4    20    3     NA        NA 0.6287402 NA
## 1    10    3     NA        NA 0.6092523 NA
```

최적조건으로 최종적합

```
# 최적 조건으로 학습나무를 학습 with 전체 데이터
tree <- randomForest(triptype ~ .,
                     data = w_trn ,
                     ntree = 20,
                     mtry = 10)

tree_all_prediction <- predict(tree, newdata = w_val)

# 최종 완성된 결정나무의 분류 성능 확인
tree_all_cm <- confusionMatrix(tree_all_prediction, w_val$triptype)
tree_all_cm
```

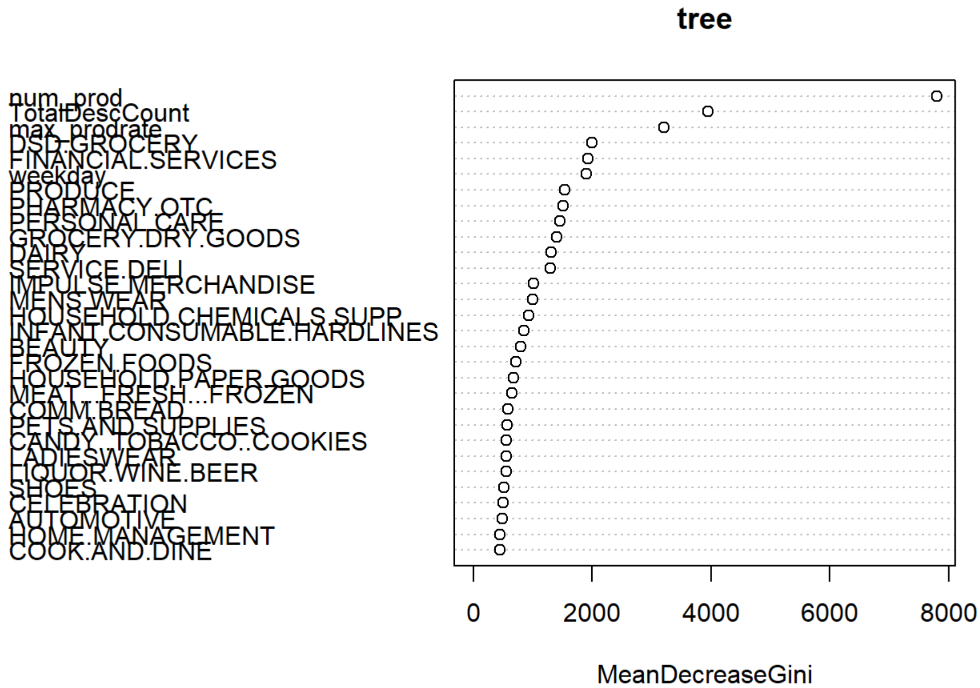
```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    3    4    5    6    7    8    9   12   14   15   18   19
##        3   1038    0    0    0    0    0   20    0    0    2    1    0
##        4      0    0    0    0    0    0    0    0    0    0    0    0
##        5      0   87  727    3    7  112   32    1    0    3    0    1
##        6      0    0    4  236    6   77    2    0    0    0    0    1
##        7      0    0   11   12 1117  143   31    3    0    6    1    1
##        8     42    3   48   34  191 3092  233    1    0   27    9    0
##        9      1    0   17    4    8   85 2261    4    0    9   35   66
##        12     0    0    1    0    0    0    0    0    0    1    0    0
##        14     0    0    0    0    0    0    0    0    0    0    0    0
##        15     0    0    1    0    2    1    2    3    0  110    1    0
##        18     0    0    0    0    0    1    3    4    0    1   51    1
##        19     0    0    0    0    0    1    9    0    0    0    0   15
##        20     0    0    1    0    0    0    2    0    0    0    4    0
##        21     0    0    2    1    1    1    5    1    2    5    4    1
##        22     2    0    0    0    0    0   19    0    0    0    0   19
##        23     0    0    0    0    0    0   10    0    0    0    0    1
##        24     6    0    3    3    7    5   20    3    0    5    6    2
##        25     2    0    7    1    5    2   41    1    0    6    7    4
##        26     0    0    0    0    0    0    1    0    0    1    0    1
##        27     1    0    0    0    1    1    1    9    0    1    0    0
##        28     0    0    0    0    0    0    2    0    0    1    0    0
##        29     0    0    0    0    0    0    5    0    0    0    3    0
##        30     1    0    1    0    2    2   17    0    0    4    0    0
##        31     0    0    0    0    0    2   62    0    0    1    0    0
##        32     0    0    0    0    4   63   40    1    0    4    5    0
##        33     0    1    2    4    4    4    1    8    0    7    2    0
##        34     0    0    2    0    1    3    1    0    0    0    1    1
##        35     1    0    7    9    8    5    2    3    0   17    7    0
##        36     1    1   34    2    2   17    3    2    0    4    5    2
##        37     0    0    5    6   64    5    2    5    0   10    1    1
##        38     1    0    4    2   41   25    3    1    0    3    1    2
##        39     2    6  100   49  208   11    1   28    0   54   19    7
##        40     0    0    6    2    5    0    0    7    0    9    2    0
##        41     0    0    0    0    0    0    1    0    0    0    1    0
##        42     0    0    1    1    0    1    2    2    0    4   12    2
##        43     0    0    0    0    0    0    0    0    0    1    1    0
##        44     0    0    0    0    0    0    0    0    0    0    1    0
##        999    0    0    3    0    0    4    3    0    0    0    0    0
##           Reference
## Prediction   20   21   22   23   24   25   26   27   28   29   30   31
##        3      2    1    1    4    7    0    4    1    0    2    2    0
##        4      0    0    0    0    0    0    0    0    0    0    1    0
##        5      4    3    5    1    7    7    3    6    2    3    7    2
##        6      0    1    1    0    0    0    0    1    0    1    0    0
##        7      1    3    4    0    9    7    4    4    4    1    5    1
```

```
##          8    2    1    4    1   14    3    3    2    1    4    2   10
##          9   51   12   90   20  108   84   23   28   50   43   84    0
##          12    0    0    0    0    0    0    0    0    0    0    0    0
##          14    0    0    0    0    0    0    0    0    0    0    0    0
##          15    0    7    2    1    5    1    1    1    1    0    1    0
##          18    0    2    0    0    1    1    1    1    2   14    1    0
##          19    0    0    9    0    0    0    0    0    0    0    0    1
##          20  106    0    2    0    4    2    4    2    0    2    2    0
##          21    0  113    3    1    6    2    1    1    2    0    2    0
##          22    0    0   63    1    1    0    1    0    0    0    1    1
##          23    0    0    3   10    0    0    0    0    0    0    0    0
##          24    1    9    7    0  368   11    7    2    5    3    5    0
##          25    2    5   10    0   17  784    2    0    2    5   36    0
##          26    0    0    1    0    4    0   48    0    0    0    1    0
##          27    1    0    1    0    2    1    1  136    1    0    0    0
##          28    1    0    0    0    0    2    1    0   56   12    1    1
##          29    0    2    0    0    1    0    0    0    5    2    0    0
##          30    1    1    2    0    2   14    1    0    1    1   96    0
##          31    0    0    8    0    0    0    0    0    0    0    0  151
##          32    0    2    2    0    3    8    0    2    1    1    3    1
##          33    0    1    2    1   15    4    1    0    1    0    1    0
##          34    1    0    2    0    4    1    0    4    1    1    4    0
##          35    3    3    4    1   12   13    1    1    2    2    7    3
##          36    6    2    9    0   22   17    4    5    4    3    7    1
##          37    1    1    0    0   13   11    2    3    3    3    3    0
##          38    2    1    4    2    6    4    4    1    0    2    5    1
##          39   17   10   30    3   97   93   11   21   15   13   27   14
##          40    1    3    3    0   34   22    3    7    0    1    4    0
##          41    0    0    0    0    0    3    0    0    1    1    0    0
##          42    5    9    1    0   17   19    4    4    4    3    9    2
##          43    0    0    1    0    3    2    1    0    3    0    0    0
##          44    0    0    0    0    2    4    0    0    0    0    0    0
##          999    0    0    0    1    3    0    0    0    0    0    0    0
##           Reference
## Prediction   32   33   34   35   36   37   38   39   40   41   42   43
##          3    1    0    0    0    1    0    0    0    0    1    0    0
##          4    0    0    0    0    0    0    0    0    0    0    0    0
##          5    4    3    1    5   36   10    4   50    2    0    3    8
##          6    0    0    3    1    0    0    1   12    0    1    0    1
##          7    3    0    5   20    3   80   46  125    3    0    1    7
##          8   10   13   34   28   35   41   10   13    0    1    5    0
##          9    5    5    0    2    4    0    0    3    0    3   10    0
##          12    0    0    0    0    1    0    0    0    0    0    1    0
##          14    0    0    0    0    0    0    0    0    0    0    0    0
##          15    0    1    0    2    1    1    0    5    0    2   18    1
##          18    3    0    0    2    1    0    0    1    0    4   10    1
##          19    0    0    0    0    0    0    1    1    0    1    0    0
##          20    0    2    1    0    1    0    0    0    0    1    7    3
##          21    1    2    0    1    3    0    0    0    1    1   12    8
##          22    0    0    0    0    1    0    0    1    0    0    4    1
##          23    0    0    0    0    0    0    0    0    0    0    0    0
##          24    7    4    2    6    4    1    0   24    0   15   46   10
##          25    8    1    0    8   13    3    4   25    4   35   98   25
##          26    0    0    0    0    2    0    0    0    0    2    6    4
##          27    1    0    2    0    3    0    2    1    0    2   14    1
##          28    0    0    0    0    1    0    0    0    0    1    7    2
##          29    0    0    0    0    0    0    0    1    0    3    3    0
##          30    3    0    0    1    2    0    0    3    0   12   10    5
##          31    0    0    0    0    0    0    0    0    0    0    0    0
##          32  416    1    2    3    4    0    6   42    2    7   26    4
##          33    3  225    1    7    4    0    5   47    1    2    6    7
##          34    1    0  113    0    0    3    1   25    5    1    1    0
##          35    3    9    3  382    2    5   26   82    2    1    8    3
##          36    7    9    5   10  550    0    4   99   11    1    8   16
##          37    2    4    3   17    6  423   18   63   40    2    1    4
##          38    6    6    5   16    6    6  335  113   24    1    4    3
##          39   68  101   34  118  167  111  313 2181   79   38   87  115
##          40   17    8    6    9   13  115   97   46 1658    5   32    2
##          41    0    0    1    0    0    0    0    0    0    3    6    0
##          42   10    1    0    2    2    0    0    3    1   22  102   12
##          43    0    0    1    1    1    2    0    3    0    3    5    9
##          44    1    0    0    0    2    0    0    4    4    6   22    1
##          999    0    0    0    2    0    0    5    0    0    0    1    0
```

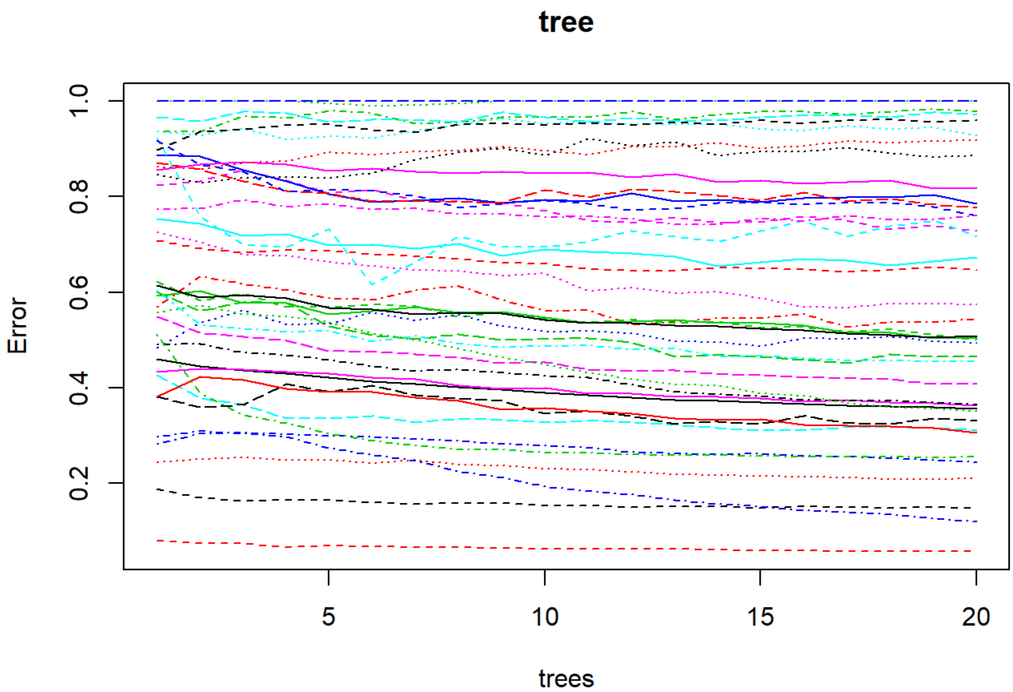
```
##           Reference
## Prediction   44  999
##        3      0  212
##        4      0    0
##        5      3   24
##        6      0    0
##        7      1   36
##        8      0  119
##        9      0  124
##        12     0    0
##        14     0    0
##        15     4    4
##        18     0    0
##        19     0    1
##        20     0    2
##        21     1    1
##        22     0   13
##        23     0    1
##        24     8   23
##        25    20    8
##        26     0    2
##        27     1    1
##        28     0    1
##        29     0    0
##        30     0    3
##        31     1    4
##        32     7    5
##        33     1    1
##        34     2    0
##        35     0    3
##        36     8   13
##        37     2    6
##        38     2   14
##        39   147   16
##        40   104    0
##        41     2    0
##        42     9    2
##        43     4    0
##        44    12    0
##        999    0 1882
##
## Overall Statistics
##
##                Accuracy : 0.6674
##                  95% CI : (0.6619, 0.6729)
##     No Information Rate : 0.1296
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.642
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: 3  Class: 4 Class: 5 Class: 6 Class: 7
## Sensitivity           0.94536 0.000e+00  0.73658 0.639566  0.66330
## Specificity           0.99036 1.000e+00  0.98355 0.995951  0.97815
## Pos Pred Value        0.79846 0.000e+00  0.61820 0.676218  0.65783
## Neg Pred Value        0.99778 9.965e-01  0.99041 0.995237  0.97866
## Prevalence            0.03883 3.466e-03  0.03491 0.013051  0.05956
## Detection Rate        0.03671 0.000e+00  0.02571 0.008347  0.03951
## Detection Prevalence  0.04598 3.537e-05  0.04159 0.012343  0.06006
## Balanced Accuracy     0.96786 5.000e-01  0.86006 0.817758  0.82073
##                      Class: 8 Class: 9 Class: 12 Class: 14 Class: 15
## Sensitivity            0.8441  0.79697 0.0000000 0.000e+00  0.371622
## Specificity            0.9616  0.96155 0.9998581 1.000e+00  0.997534
## Pos Pred Value         0.7661  0.69805 0.0000000       NaN  0.614525
## Neg Pred Value         0.9764  0.97699 0.9969225 9.999e-01  0.993380
## Prevalence             0.1296  0.10034 0.0030770 7.074e-05  0.010469
## Detection Rate         0.1094  0.07997 0.0000000 0.000e+00  0.003891
## Detection Prevalence   0.1427  0.11456 0.0001415 0.000e+00  0.006331
## Balanced Accuracy      0.9029  0.87926 0.4999290 5.000e-01  0.684578
##                      Class: 18 Class: 19 Class: 20 Class: 21 Class: 22
## Sensitivity           0.283333 0.1171875  0.509615  0.588542  0.229927
```

```
## Sensitivity          0.283333 0.1171875  0.509615  0.588542  0.229927
## Specificity          0.998042 0.9991473  0.998504  0.997436  0.997679
## Pos Pred Value       0.481132 0.3846154  0.716216  0.610811  0.492187
## Neg Pred Value       0.995420 0.9959979  0.996373  0.997188  0.992503
## Prevalence           0.006366 0.0045271  0.007357  0.006791  0.009691
## Detection Rate       0.001804 0.0005305  0.003749  0.003997  0.002228
## Detection Prevalence  0.003749 0.0013794  0.005234  0.006543  0.004527
## Balanced Accuracy     0.640688 0.5581674  0.754059  0.792989  0.613803
##                     Class: 23 Class: 24 Class: 25 Class: 26 Class: 27
## Sensitivity          0.2127660   0.46760   0.70000  0.352941  0.583691
## Specificity          0.9994686   0.99054   0.98501  0.999112  0.998253
## Pos Pred Value       0.4000000   0.58599   0.65827  0.657534  0.735135
## Neg Pred Value       0.9986902   0.98484   0.98759  0.996880  0.996547
## Prevalence           0.0016623   0.02783   0.03961  0.004810  0.008241
## Detection Rate       0.0003537   0.01302   0.02773  0.001698  0.004810
## Detection Prevalence 0.0008842   0.02221   0.04212  0.002582  0.006543
## Balanced Accuracy    0.6061173   0.72907   0.84251  0.676026  0.790972
##                     Class: 28 Class: 29 Class: 30 Class: 31 Class: 32
## Sensitivity           0.335329 1.626e-02  0.302839  0.798942   0.71724
## Specificity           0.998826 9.992e-01  0.996817  0.997223   0.99101
## Pos Pred Value        0.629213 8.000e-02  0.518919  0.659389   0.62556
## Neg Pred Value        0.996062 9.957e-01  0.992132  0.998645   0.99406
## Prevalence            0.005906 4.350e-03  0.011212  0.006685   0.02051
## Detection Rate        0.001981 7.074e-05  0.003395  0.005341   0.01471
## Detection Prevalence  0.003148 8.842e-04  0.006543  0.008099   0.02352
## Balanced Accuracy     0.667078 5.077e-01  0.649828  0.898082   0.85413
##                     Class: 33 Class: 34 Class: 35 Class: 36 Class: 37
## Sensitivity           0.569620  0.509009   0.59409   0.63291   0.52809
## Specificity           0.994835  0.997647   0.99066   0.98745   0.98883
## Pos Pred Value        0.609756  0.631285   0.59688   0.61521   0.57945
## Neg Pred Value        0.993908  0.996120   0.99056   0.98835   0.98628
## Prevalence            0.013970  0.007852   0.02274   0.03073   0.02833
## Detection Rate        0.007958  0.003997   0.01351   0.01945   0.01496
## Detection Prevalence  0.013051  0.006331   0.02264   0.03162   0.02582
## Balanced Accuracy     0.782228  0.753328   0.79238   0.81018   0.75846
##                     Class: 38 Class: 39 Class: 40 Class: 41 Class: 42
## Sensitivity            0.38155   0.73360   0.90256 0.0169492  0.180851
## Specificity            0.98828   0.91186   0.97870 0.9994305  0.994009
## Pos Pred Value         0.51067   0.49445   0.74651 0.1578947  0.380597
## Neg Pred Value         0.98034   0.96681   0.99313 0.9938418  0.983504
## Prevalence             0.03105   0.10515   0.06497 0.0062602  0.019948
## Detection Rate         0.01185   0.07714   0.05864 0.0001061  0.003608
## Detection Prevalence   0.02320   0.15601   0.07855 0.0006720  0.009479
## Balanced Accuracy      0.68492   0.82273   0.94063 0.5081898  0.587430
##                     Class: 43 Class: 44 Class: 999
## Sensitivity          0.0355731 0.0353982    0.74653
## Specificity          0.9988580 0.9983175    0.99915
## Pos Pred Value       0.2195122 0.2033898    0.98845
## Neg Pred Value       0.9913576 0.9884104    0.97577
## Prevalence           0.0089482 0.0119898    0.08916
## Detection Rate       0.0003183 0.0004244    0.06656
## Detection Prevalence 0.0014501 0.0020867    0.06734
## Balanced Accuracy    0.5172156 0.5168579    0.87284
```

```
# 시각화
varImpPlot(tree)
```




```
plot(tree)
```



5. 설정 모델로 예측 및 평가