

Ödev 2: Decision Tree and Random Forest

Adı Soyadı : Aysu İğdi

Numara : 200207016

GitHub [repository linki](#)

Decision Tree and Random Forest

Decision Trees

RStudio interface showing the 'wholesale_customers_data' dataset loaded. The console displays the data structure:

```
> View(wholesale_customers_data)
> str(wholesale_customers_data)
'spc_tbl' [440 x 8] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ Channel      : num [1:440] 2 2 2 1 2 2 2 2 1 2 ...
 $ Region       : num [1:440] 3 3 3 3 3 3 3 3 3 3 ...
 $ Fresh        : num [1:440] 12669 7057 6353 13265 22615 ...
 $ Milk         : num [1:440] 9656 9810 8808 1196 5410 ...
 $ Grocery      : num [1:440] 7561 9568 7684 4221 7198 ...
 $ Frozen       : num [1:440] 214 1762 2405 6404 3915 ...
 $ Detergents_Paper: num [1:440] 2674 3293 3516 507 1777 ...
 $ Delicassen   : num [1:440] 1338 1776 7844 1788 5185 ...
```

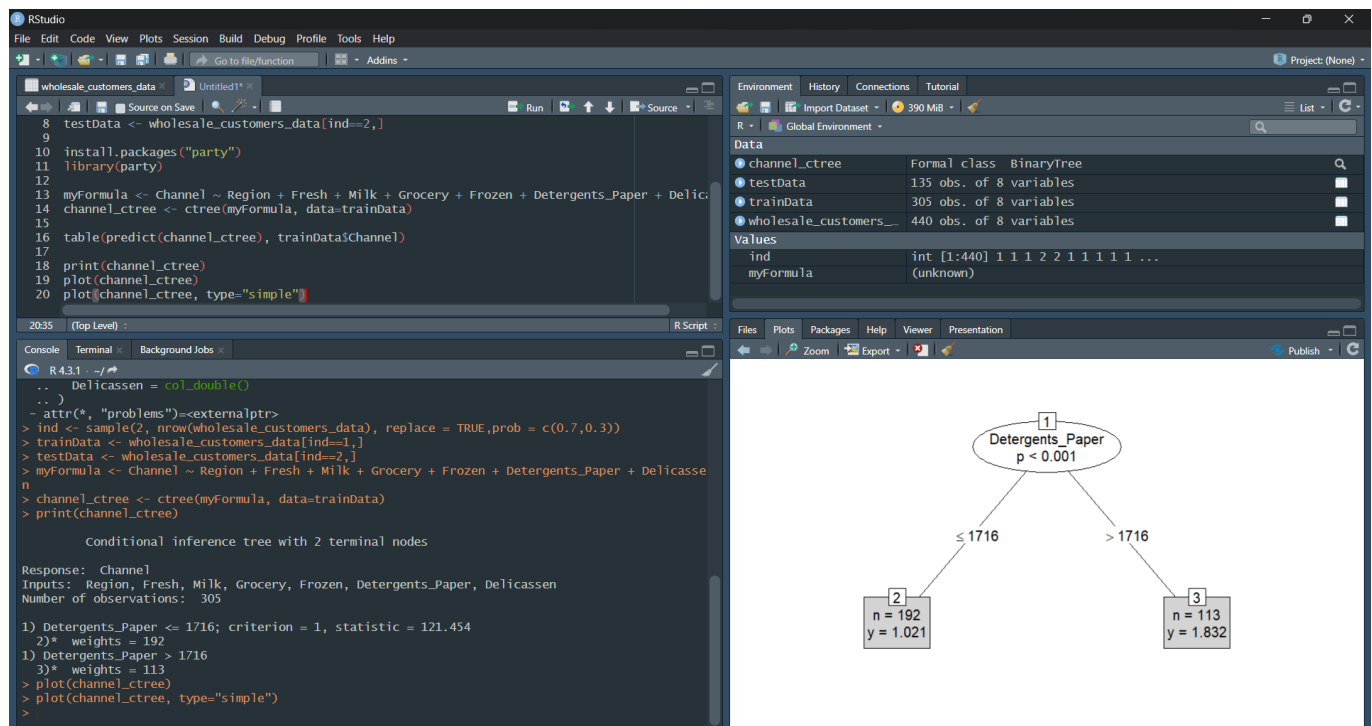
The Environment pane shows the dataset 'wholesale_customers_data' with 440 obs. of 8 variables.

RStudio interface showing the 'wholesale_customers_data' dataset being split into training and testing sets. The console displays the code:

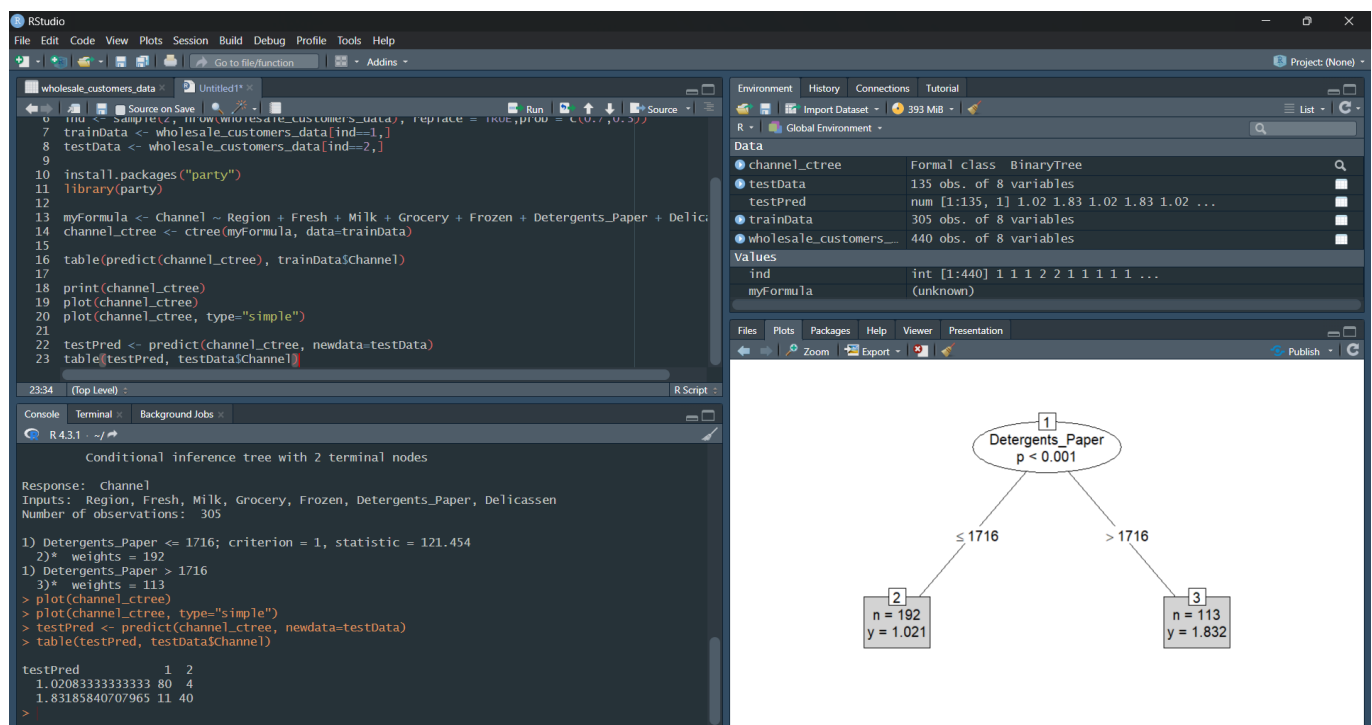
```
1 View(wholesale_customers_data)
2
3 str(wholesale_customers_data)
4
5 set.seed(1234)
6 ind <- sample(2, nrow(wholesale_customers_data), replace = TRUE, prob = c(0.7, 0.3))
7 trainData <- wholesale_customers_data[ind==1,]
8 testData <- wholesale_customers_data[ind==2,]
9
10
```

The Environment pane shows the resulting datasets: 'trainData' (305 obs. of 8 variables) and 'testData' (135 obs. of 8 variables). The 'wholesale_customers_data' dataset remains in the environment.

trainData

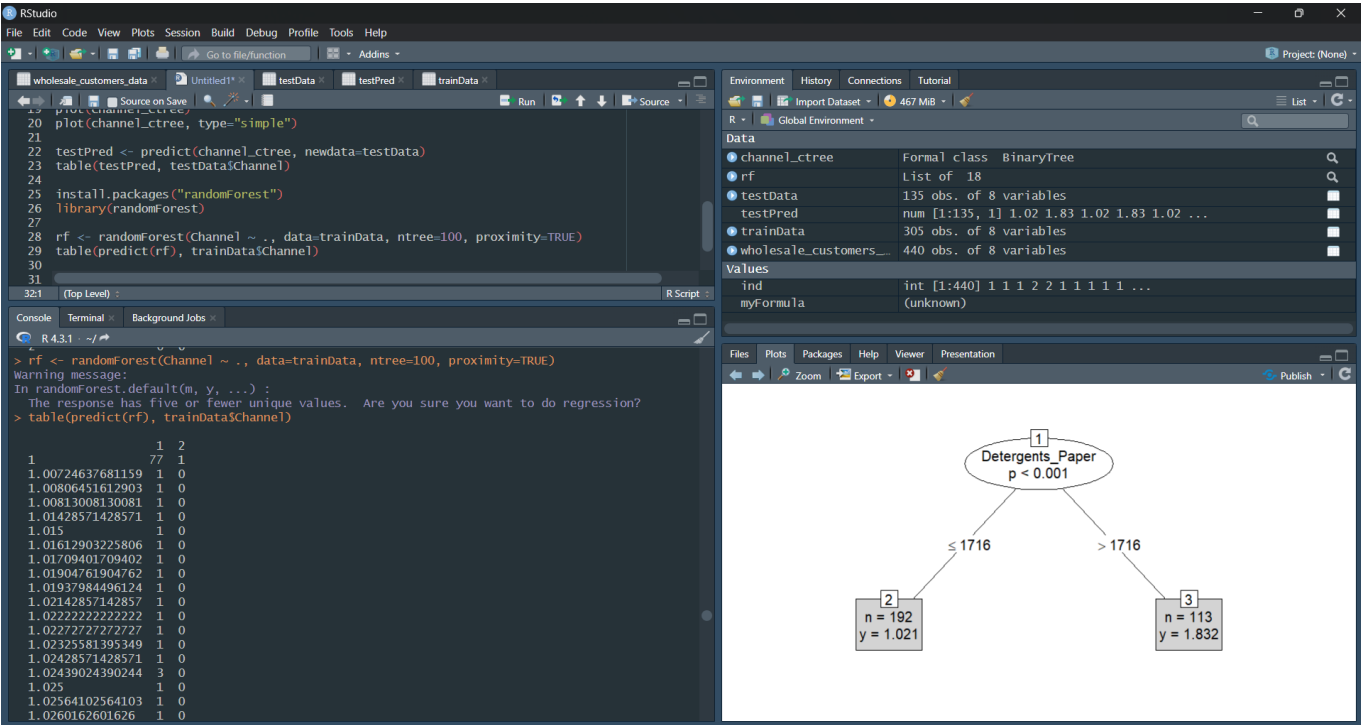


predict with testData



testData'da 91 adet **channel=1** değerinin 80 adetini doğru, 11 adetini yanlış; 44 adet **channel=2** değerinin 40 adetini doğru, 4 adetini yanlış tahmin etti.

Random Forest



attributes

