

RStudio interface showing R code, console output, and environment pane.

Code Editor:

```
1 #PROJECT : BIKE RENTAL PREDICTION ...
2 #P : TO PREDICT THE COUNT OF BIKE RENTAL
3 #S : TO BUILD RANDOM FOREST Alogrithm
4
5 #Importing Library
6 library(randomForest)
7 library(readxl)
8
9 #Importing file & Initial Aolyssis
10 day1 <- read_excel("day1.xlsx")
11 day1
12
13 head(day1)
14 tail(day1)
15 names(day1)
16 str(day1)
17 summary(day1)
18
```

Console:

```
R 4.0.2 - 731 x 12
# A tibble: 731 x 12
  season  yr  mnth holiday weekday workingday weathersit temp atemp hum windspeed cnt
  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1      1      0      1      0      6      0      2  0.344 0.364 0.806  0.160  985
2      1      0      1      0      0      0      2  0.363 0.354 0.696  0.249  801
3      1      0      1      0      1      1      1  0.196 0.189 0.437  0.248 1349
4      1      0      1      0      2      1      1  0.2    0.212 0.590  0.160 1562
5      1      0      1      0      3      1      1  0.227 0.229 0.437  0.187 1600
6      1      0      1      0      4      1      1  0.204 0.233 0.518  0.0806 1606
7      1      0      1      0      5      1      2  0.197 0.209 0.499  0.169 1510
8      1      0      1      0      6      0      2  0.165 0.162 0.536  0.267  959
9      1      0      1      0      0      0      1  0.138 0.116 0.434  0.362  822
10     1      0      1      0      1      1      1  0.151 0.151 0.483  0.223 1321
# ... with 721 more rows
>
```

Environment:

Global Environment

day1 731 obs. of 12 variables

Files:

Name	Size	Modified
day1.xlsx	79 KB	May 7, 2024, 8:33 PM
day1_forest_output.csv	6 KB	May 8, 2024, 4:53 PM
day1_liner_output.csv	16.4 KB	May 8, 2024, 12:15 PM
day1.xlsx	58.7 KB	May 8, 2024, 12:01 PM
dy1_forest.csv	58.9 KB	May 8, 2024, 9:42 PM
groceries.csv	489.1 KB	May 5, 2024, 9:56 PM
liner_output.csv	3.4 KB	Apr 27, 2024, 11:16 PM
R		
r file.R	416 B	Apr 6, 2024, 8:35 PM
regression data.csv	6.3 KB	Apr 27, 2024, 10:12 PM
Remission.csv	2.9 KB	Apr 28, 2024, 7:43 PM
spss dataset.csv	3.7 MB	Apr 6, 2024, 7:32 PM
survey_data.csv	20.1 KB	Apr 27, 2024, 7:43 PM
Bike rental Prediction.R	1.7 KB	May 9, 2024, 12:23 PM

RStudio interface showing R code for data analysis and environment details.

Code Editor:

```
9 #Importing file & Initial Analysis
10 day1 <- read_excel("day1.xlsx")
11 day1
12
13 head(day1)
14 tail(day1)
15 names(day1)
16 str(day1)
17 summary(day1)
18
19 #Convert categorical variables to factors
20 day1$season <- as.factor(day1$season)
21 day1$yr <- as.factor(day1$yr)
22 day1$mnth <- as.factor(day1$mnth)
23 day1$holiday <- as.factor(day1$holiday)
24 day1$weekday <- as.factor(day1$weekday)
25 day1$weathersit <- as.factor(day1$weathersit)
26
```

Console:

```
R 4.0.2 - /
# - with 721 more rows
> summary(day1)
   season      yr      mnth    holiday    weekday    workingday
Min.   :1.000  Min.   :0.0000  Min.   :1.00  Min.   :0.00000  Min.   :0.000  Min.   :0.000
1st Qu.:1.000  1st Qu.:0.0000  1st Qu.:4.00  1st Qu.:0.00000  1st Qu.:1.000  1st Qu.:0.000
Median :1.000  Median :1.0000  Median :7.00  Median :0.00000  Median :3.000  Median :1.000
Mean   :2.497  Mean   :0.5007  Mean   :6.52  Mean   :0.02873  Mean   :2.997  Mean   :0.684
3rd Qu.:3.000  3rd Qu.:1.0000  3rd Qu.:10.00  3rd Qu.:0.00000  3rd Qu.:5.000  3rd Qu.:1.000
Max.   :4.000  Max.   :1.0000  Max.   :12.00  Max.   :1.00000  Max.   :6.000  Max.   :1.000

 weathersit    temp      atemp      hum    windspeed      cnt
Min.   :1.000  Min.   :0.05913  Min.   :0.07907  Min.   :0.0000  Min.   :0.02239  Min.   : 22
1st Qu.:1.000  1st Qu.:0.33708  1st Qu.:0.33784  1st Qu.:0.5200  1st Qu.:0.13495  1st Qu.:3152
Median :1.000  Median :0.49833  Median :0.48673  Median :0.6267  Median :0.18097  Median :4548
Mean   :1.395  Mean   :0.49538  Mean   :0.47435  Mean   :0.6279  Mean   :0.19049  Mean   :4504
3rd Qu.:2.000  3rd Qu.:0.65542  3rd Qu.:0.68860  3rd Qu.:0.7502  3rd Qu.:0.23321  3rd Qu.:5956
```

Environment:

Global Environment

Data

day1 731 obs. of 12 variables

Files:

Name	Size	Modified
day.xlsx	79 KB	May 7, 2024, 8:33 PM
day1_forest_output.csv	6 KB	May 8, 2024, 4:53 PM
day1_liner_output.csv	16.4 KB	May 8, 2024, 12:15 PM
day1.xlsx	58.7 KB	May 8, 2024, 12:01 PM
day1_forest.csv	58.9 KB	May 8, 2024, 9:42 PM
groceries.csv	489.1 KB	May 5, 2024, 9:56 PM
liner_output.csv	3.4 KB	Apr 27, 2024, 11:16 PM
R		
r file.R	416 B	Apr 6, 2024, 8:35 PM
regression data.csv	6.3 KB	Apr 27, 2024, 10:12 PM
Remission.csv	2.9 KB	Apr 28, 2024, 7:43 PM
spss dataset.csv	3.7 MB	Apr 6, 2024, 7:32 PM
survey_data.csv	20.1 KB	Apr 27, 2024, 7:43 PM
Bike rental Prediction.R	1.7 KB	May 9, 2024, 12:23 PM

rstudio

Used 19.3 of 50 hours in May, 2024

▶ Start Lab ■ End Lab ✕

PG R Studio interface showing the R script editor, Environment pane, and Files pane.

Code Editor:

```
13 head(day1)
14 tail(day1)
15 names(day1)
16 str(day1)
17 summary(day1)
18
19 #Convert categorical variables to factors
20 day1$season <- as.factor(day1$season)
21 day1$yr <- as.factor(day1$yr)
22 day1$mnth <- as.factor(day1$mnth)
23 day1$holiday <- as.factor(day1$holiday)
24 day1$weekday <- as.factor(day1$weekday)
25 day1$weathersit <- as.factor(day1$weathersit)
26
27
28 #splitting data into training and testing set
29 set.seed(1)
30 indexes <- sample(1:nrow(day1), size = 0.7 * nrow(day1))
31
32 (Top Level) :
```

Environment Pane:

Global Environment

day1 731 obs. of 12 variables

Files Pane:

Name	Size	Modified
day.xlsx	79 KB	May 7, 2024, 8:33 PM
day1_forest_output.csv	6 KB	May 8, 2024, 4:53 PM
day1_liner_output.csv	16.4 KB	May 8, 2024, 12:15 PM
day1.xlsx	58.7 KB	May 8, 2024, 12:01 PM
day1_forest.csv	58.9 KB	May 8, 2024, 9:42 PM
groceries.csv	489.1 KB	May 5, 2024, 9:56 PM
liner_output.csv	3.4 KB	Apr 27, 2024, 11:16 PM
R		
r file.R	416 B	Apr 6, 2024, 8:35 PM
regression data.csv	6.3 KB	Apr 27, 2024, 10:12 PM
Remission.csv	2.9 KB	Apr 28, 2024, 7:43 PM
sps dataset.csv	3.7 MB	Apr 6, 2024, 7:32 PM
survey_data.csv	20.1 KB	Apr 27, 2024, 7:43 PM
Bike rental Prediction R	1.7 KB	May 9, 2024, 12:23 PM

R Studio interface showing R code for data preprocessing and model training.

```
21 day1$yr <- as.factor(day1$yr)
22 day1$mnth <- as.factor(day1$mnth)
23 day1$holiday <- as.factor(day1$holiday)
24 day1$weekday <- as.factor(day1$weekday)
25 day1$weathersit <- as.factor(day1$weathersit)
26
27
28 #Splitting data into training and testing set
29 set.seed(1)
30 indexes <- sample(1:nrow(day1), size = 0.7 * nrow(day1))
31 train_data <- day1[indexes, ]
32 test_data <- day1[-indexes, ]
33
34 #Creating Random Forest Model
35 rf <- randomForest(cnt ~ ., data = train_data, ntree = 500)
36 rf
37
38 #Making Prediction
39 predict(rf, test_data)
```

Console output:

```
R 4.0.2 ~ /
> day1$season <- as.factor(day1$season)
> day1$yr <- as.factor(day1$yr)
> day1$mnth <- as.factor(day1$mnth)
> day1$holiday <- as.factor(day1$holiday)
> day1$weekday <- as.factor(day1$weekday)
> day1$weathersit <- as.factor(day1$weathersit)
> set.seed(1)
> indexes <- sample(1:nrow(day1), size = 0.7 * nrow(day1))
> train_data <- day1[indexes, ]
> test_data <- day1[-indexes, ]
>
```

Environment pane:

Object	Class	Attributes
day1	data.frame	731 obs. of 12 variables
test_data	data.frame	220 obs. of 12 variables
train_data	data.frame	511 obs. of 12 variables
indexes	integer	[1:511] 679 129 509 471 299 270 187 307 597 277 ...

Files pane:

Name	Size	Modified
day.xlsx	79 KB	May 7, 2024, 8:33 PM
day1_forest_output.csv	6 KB	May 8, 2024, 4:53 PM
day1_liner_output.csv	16.4 KB	May 8, 2024, 12:15 PM
day1.xlsx	58.7 KB	May 8, 2024, 12:01 PM
day1_forest.csv	58.9 KB	May 8, 2024, 9:42 PM
groceries.csv	489.1 KB	May 5, 2024, 9:56 PM
liner_output.csv	3.4 KB	Apr 27, 2024, 11:16 PM
R		
r file.R	416 B	Apr 6, 2024, 8:35 PM
regression data.csv	6.3 KB	Apr 27, 2024, 10:12 PM
Remission.csv	2.9 KB	Apr 28, 2024, 7:43 PM
spss dataset.csv	3.7 MB	Apr 6, 2024, 7:32 PM
survey_data.csv	20.1 KB	Apr 27, 2024, 7:43 PM
Bike rental Prediction.R	1.7 KB	May 9, 2024, 12:23 PM

File

Edit

Code

View

Plots

Session

Build

Debug

Profile

Tools

Help

labruser

Project: (Name)

Go to file/function

Addins

on

Untitled3*

Untitled4*

Untitled5*

Untitled6*

Untitled7*

Untitled8*

Bike rental Prediction R

Source on Save

Run

Source

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

36:1

(Top Level) :

day1\$weathersit <- as.factor(day1\$weathersit)

#Splitting data into training and testing set

set.seed(1)

indexes <- sample(1:nrow(day1), size = 0.7 * nrow(day1))

train_data <- day1[indexes,]

test_data <- day1[-indexes,]

#Creating Random Forest Model

rf <- randomForest(cnt ~ ., data = train_data, ntree = 500)

#Making Prediction

predictions <- predict(rf, newdata = test_data)

#Evaluation

actual <- test_data\$cnt

36:1

(Top Level) :

Console

Terminal

Jobs

R 4.0.2 - /

> train_data <- day1[indexes,]

> test_data <- day1[-indexes,]

> rf <- randomForest(cnt ~ ., data = train_data, ntree = 500)

> day1\$weathersit <- as.factor(day1\$weathersit)

> rf

Call:

randomForest(formula = cnt ~ ., data = train_data, ntree = 500)

Type of random forest: regression

Number of trees: 500

No. of variables tried at each split: 3

Mean of squared residuals: 532261

% Var explained: 85.37

> |

Environment

History

Connections

Tutorial

Import Dataset

197 MB

Global Environment

Data

day1

731 obs. of 12 variables

rf

List of 18

test_data

220 obs. of 12 variables

train_data

511 obs. of 12 variables

Values

Files

Plots

Packages

Help

Viewer

New Folder

New Blank File

Upload

Delete

Rename

More

Home

Name

Size

Modified

day.xlsx

79 KB

May 7, 2024, 8:33 PM

day1_forest_output.csv

6 KB

May 8, 2024, 4:53 PM

day1_liner_output.csv

16.4 KB

May 8, 2024, 12:15 PM

day1.xlsx

58.7 KB

May 8, 2024, 12:01 PM

dyl_forest.csv

58.9 KB

May 8, 2024, 9:42 PM

groceries.csv

489.1 KB

May 5, 2024, 9:56 PM

liner_output.csv

3.4 KB

Apr 27, 2024, 11:16 PM

R

416 B

Apr 6, 2024, 8:35 PM

r file.R

6.3 KB

Apr 27, 2024, 10:12 PM

regression data.csv

2.9 KB

Apr 28, 2024, 7:43 PM

Remission.csv

3.7 MB

Apr 6, 2024, 7:32 PM

spss dataset.csv

20.1 KB

Apr 27, 2024, 7:43 PM

survey_data.csv

1.7 KB

May 9, 2024, 12:23 PM

Bike rental Prediction.R


```
Console Terminal Jobs
R 4.0.2 .- /
Call:
  randomForest(formula = cnt ~ ., data = train_data, ntree = 500)
    Type of random forest: regression
    Number of trees: 500
    No. of variables tried at each split: 3

    Mean of squared residuals: 532261
    % Var explained: 85.37

> predictions <- predict(rf, newdata = test_data)
> actual <- test_data$cnt
> mse <- mean((predictions - actual)^2) # Mean Squared Error
[1] 453616.4
> |
```

The screenshot displays the RStudio interface. The top pane shows the Environment window with the following data:

Variable	Value
actual	num [1:220] 1321 1162 1486 1421 1688 ...
indexes	int [1:511] 679 129 589 471 299 278 187 387 597 277 ...
mise	453616.41285222
predictions	Named num [1:220] 1486 1547 1469 1490 1440 ...

The bottom pane shows the File Explorer window with the following files and folders:

Name	Size	Modified
day.xlsx	79 KB	May 7, 2024, 8:33 PM
day1_forest_output.csv	6 KB	May 8, 2024, 4:53 PM
day1_liner_output.csv	16.4 KB	May 8, 2024, 12:15 PM
day1.xlsx	58.7 KB	May 8, 2024, 12:01 PM
day1_forest.csv	58.9 KB	May 8, 2024, 9:42 PM
groceries.csv	489.1 KB	May 5, 2024, 9:55 PM
liner_output.csv	3.4 KB	Apr 27, 2024, 11:16 PM
R		
r file.R	416 B	Apr 6, 2024, 8:35 PM
regression data.csv	6.3 KB	Apr 27, 2024, 10:12 PM
Remission.csv	2.9 KB	Apr 28, 2024, 7:43 PM
spss dataset.csv	3.7 MB	Apr 6, 2024, 7:32 PM
survey_data.csv	20.1 KB	Apr 27, 2024, 7:43 PM
Bike rental Prediction.R	1.7 KB	May 9, 2024, 12:23 PM

Actual vs. Predicted Count



