VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Lab Assignment no: 5
Decision Tree

Submitted by

Name of the Student: S. MANOJ RATHINAM

Reg no: 20MIS1010

Course Program: Int. M.Tech

Course code: ECE3502

Course Title: IOT Domain Analyst

Batch: 2020-2025

Submitted to

Dr. Rekha D

Submitted on

29-03-2024

Decision Tree

1. For the Given Indian diabetes find the Decision Tree.

Code:

```
library(rpart)

library(rpart.plot)

data <- read.csv("C:/Users/Student/Downloads/indians-diabetes.csv")

data

set.seed(123)

train_index<-sample(1:nrow(data),0.7*nrow(data),replace=FALSE)

train_data<-data[train_index,]

test_data<-data[-train_index,]

d_model<-rpart(label~.,data =train_data, method ="class",parms
=list(split="information"))

rpart.plot(d_model,type=4,extra =101)


predictions <- predict(d_model,test_data,type="class")

confusion_matrix <-  table(prediction = (predictions), actual = test_data$label)

print(confusion_matrix)

correct_prediction = confusion_matrix[1,1] + confusion_matrix[2,2]

total_prediction =
confusion_matrix[1,1]+confusion_matrix[1,2]+confusion_matrix[2,1]+confusio
n_matrix[2,2]

acc = correct_prediction / total_prediction

print(acc*100)
```

Output:

```
> data <- read.csv("C:/Users/Student/Downloads/indians-diabetes.csv")
> set.seed(123)
> library(rpart)
> library(rpart)
> train_index<-sample(1:nrow(data),0.7*nrow(data),replace=FALSE)
> train_data<-data[train_index,]
> test_data<-data[-train_data,]
Error in xj[i] : invalid subscript type 'list'
> test_data<-data[-train_index,]
> library(rpart.plot)
Warning message:
package ‘rpart.plot’ was built under R version 4.2.3
>
> d_model<-rpart(label~.,data =train_data, method ="class",parms
=list(split="information"))
> rpart.plot(d_model,type=4,extra =101)
> predictions <- predict(d_model,test_data,type="class")
> confusion_matrix <-  table(prediction = (predictions), actual =
test_data$label)
> print(confusion_matrix)
          actual
prediction   0   1
```

```
    0 123  37
    1  27  44
> correct_prediction = confusion_matrix[1,1] + confusion_matrix[2,2]
> total_prediction =
confusion_matrix[1,1]+confusion_matrix[1,2]+confusion_matrix[2,1]+confusio
n_matrix[2,2]
> acc = correct_prediction / total_prediction
> acc
[1] 0.7229437
> print(acc*100)
[1] 72.29437
>
```

Screenshots:

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins

Source

Console | Terminal | Background Jobs

R 4.2.2 · ~/

```
> data <- read.csv("C:/Users/Student/Downloads/indians-diabetes.csv")
> data
    glucose  bp skin insulin  bmi pedigree age label
1       148  72   35       0 33.6    0.627  50     1
2        85  66   29       0 26.6    0.351  31     0
3       183  64    0       0 23.3    0.672  32     1
4        89  66   23      94 28.1    0.167  21     0
5       137  40   35     168 43.1    2.288  33     1
6       116  74    0       0 25.6    0.201  30     0
7        78  50   32      88 31.0    0.248  26     1
8       115   0    0       0 35.3    0.134  29     0
9       197  70   45     543 30.5    0.158  53     1
10      125  96    0       0  0.0    0.232  54     1
11      110  92    0       0 37.6    0.191  30     0
12      168  74    0       0 38.0    0.537  34     1
13      139  80    0       0 27.1    1.441  57     0
14      189  60   23     846 30.1    0.398  59     1
15      166  72   19     175 25.8    0.587  51     1
16      100   0    0       0 30.0    0.484  32     1
17      118  84   47     230 45.8    0.551  31     1
18      107  74    0       0 29.6    0.254  31     1
19      103  30   38      83 43.3    0.183  33     0
20      115  70   30      96 34.6    0.529  32     1
21      126  88   41     235 39.3    0.704  27     0
22       99  84    0       0 35.4    0.388  50     0
23      196  90    0       0 39.8    0.451  41     1
24      119  80   35       0 29.0    0.263  29     1
25      143  94   33     146 36.6    0.254  51     1
26      125  70   26     115 31.1    0.205  41     1
27      147  76    0       0 39.4    0.257  43     1
28       97  66   15     140 23.2    0.487  22     0
29      145  82   19     110 22.2    0.245  57     0
30      117  92    0       0 34.1    0.337  38     0
31      109  75   26       0 36.0    0.546  60     0
32      158  76   36     245 31.6    0.851  28     1
33       88  58   11      54 24.8    0.267  22     0
34       92  92    0       0 19.9    0.188  28     0
35      122  78   31       0 27.6    0.512  45     0
36      103  60   33     192 24.0    0.966  33     0
37      138  76    0       0 33.2    0.420  35     0
38      102  76   37       0 32.9    0.665  46     1
39       90  68   42       0 38.2    0.503  27     1
40      111  72   47     207 37.1    1.390  56     1
41      180  64   25      70 34.0    0.271  26     0
42      133  84    0       0 40.2    0.696  37     0
43      106  92   18       0 22.7    0.235  48     0
44      171 110   24     240 45.4    0.721  54     1
45      159  64    0       0 27.4    0.294  40     0
46      180  66   39       0 42.0    1.893  25     1
47      146  56    0       0 29.7    0.564  29     0
48       71  70   27       0 28.0    0.586  22     0
49      103  66   32       0 39.1    0.344  31     1
50      105   0    0       0  0.0    0.305  24     0
51      103  80   11      82 19.4    0.491  22     0
52      101  50   15      36 24.2    0.526  26     0
53       88  66   21      23 24.4    0.342  30     0
54      176  90   34     300 33.7    0.467  58     1
55      150  66   42     342 34.7    0.718  42     0
```

VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins

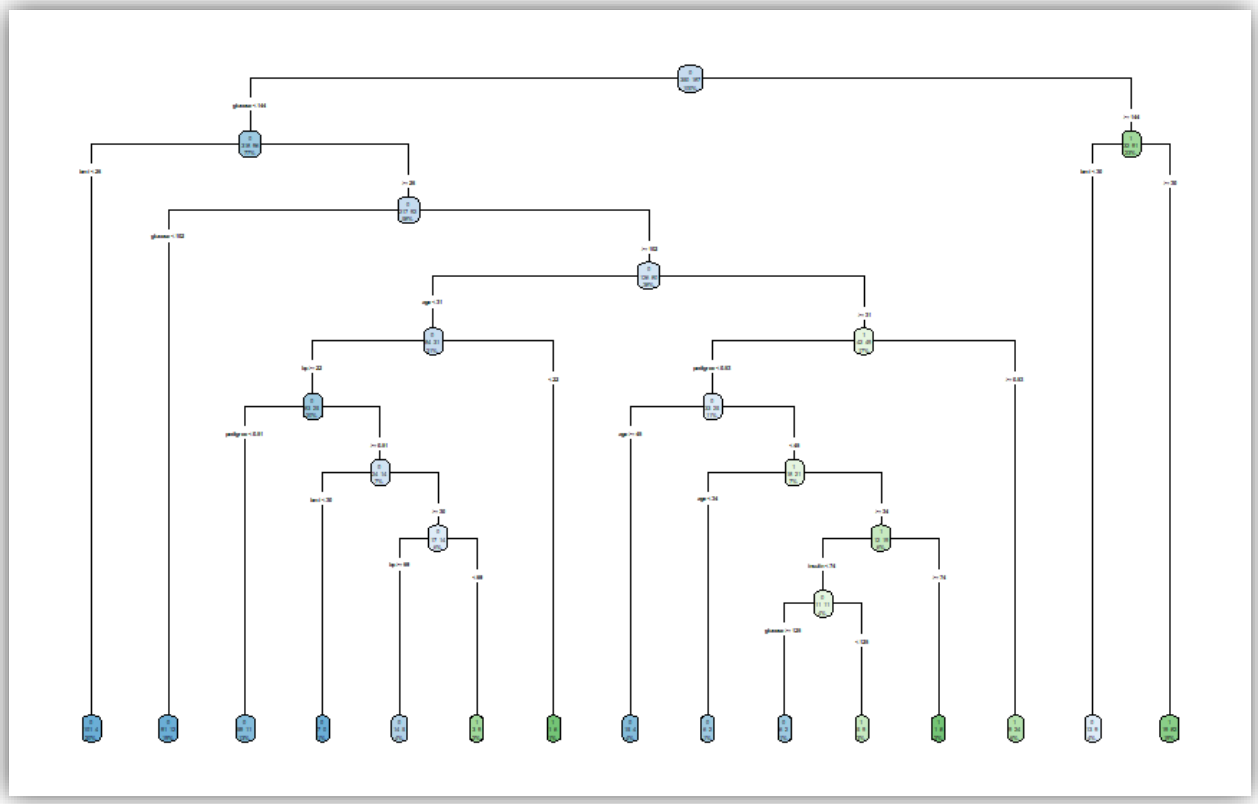
Source

Console | Terminal | Background Jobs

R 4.2.2 · ~/

```
 [ reached 'max' / getOption("max.print") -- omitted 643 rows ]
> set.seed(123)
> library(rpart)
> library(rpart)
> train_index<-sample(1:nrow(data),0.7*nrow(data),replace=FALSE)
> train_data<-data[train_index,]
> test_data<-data[-train_data,]
Error in xj[i] : invalid subscript type 'list'
> test_data<-data[-train_index,]
> d_model<-rpart(label~.,data =train_Data, method ="class",parms =list(split="information")
+                rpart.plot(diabetes_model,type=4,extra =101)
Error: unexpected symbol in:
"d_model<-rpart(label~.,data =train_Data, method ="class",parms =list(split="information")
               rpart.plot"
> rpart.plot(diabetes_model,type=4,extra =101)
Error in rpart.plot(diabetes_model, type = 4, extra = 101) :
  could not find function "rpart.plot"
> library(rpart.plot)
Warning message:
package 'rpart.plot' was built under R version 4.2.3
> d_model<-rpart(label~.,data =train_Data, method ="class",parms =list(split="information")
+ rpart.plot(diabetes_model,type=4,extra =101)
Error: unexpected symbol in:
"d_model<-rpart(label~.,data =train_Data, method ="class",parms =list(split="information")
rpart.plot"
>
> d_model<-rpart(label~.,data =train_Data, method ="class",parms =list(split="information"))
Error in is.data.frame(data) : object 'train_Data' not found
> d_model<-rpart(label~.,data =train_data, method ="class",parms =list(split="information"))
> rpart.plot(diabetes_model,type=4,extra =101)
Error in rpart.plot(diabetes_model, type = 4, extra = 101) :
  object 'diabetes_model' not found
> rpart.plot(d_model,type=4,extra =101)
> predictions <- predict(d_model,test_data,type="class")
> confusion_matrix <-  table(prediction = max.col(predictions), actual = test_data$label)
> print(confusion_matrix)
          actual
prediction   0   1
         1 150  81
> confusion_matrix <-  table(prediction = (predictions), actual = test_data$label)
> print(confusion_matrix)
          actual
prediction   0   1
         0 123  37
         1  27  44
```

```
> correct_prediction = confusion_matrix[1,1] + confusion_matrix[2,2]
> total_prediction = confusion_matrix[1,1]+confusion_matrix[1,2]+confusion_matrix[2,1]+confusion_matrix[2,2]
> acc = correct_prediction / total_prediction
> acc
[1] 0.7229437
> print(acc*100)
[1] 72.29437
>
```



Decision Tree