MSMS 206 : Practical 09

Ananda Biswas

May 6, 2025

🎯 **Question :** Fit a decision tree model using the **Loan Defaulters' Dataset** given as follows :

Home Owner	Married Status	Defaulted	Annual Income($)
yes	single	no	125000
no	married	no	100000
no	single	no	70000
yes	married	no	120000
no	divorcee	yes	95000
no	married	no	60000
yes	divorcee	no	220000
no	single	yes	85000
no	married	no	75000
no	single	yes	90000

Will a divorcee home owner with annual income $120000 default in his loan ?

➔ Building Decision Tree Model

```
library(rpart)
library(rpart.plot)
```

```
df <- read.csv('https://raw.githubusercontent.com/sakunisgithub/data_sets/refs/heads/master/msc_semester_2/loan_defaulter_data.csv')
```

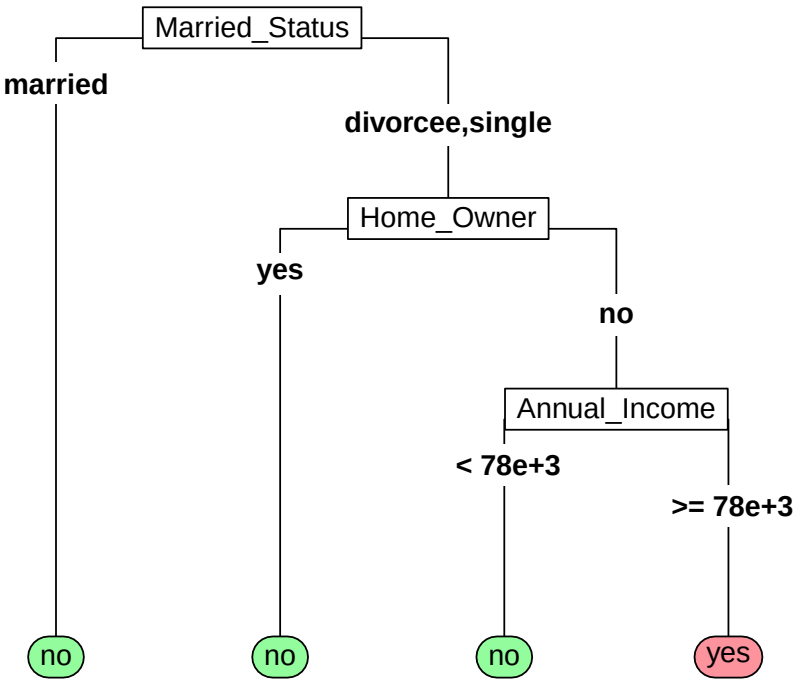
```r
tree_model <- rpart(Defaulted ~ .,
                    data = df,
                    method = "class",
                    parms = list(split = "information"),
                    control = rpart.control(minsplit = 2,
                                            minbucket = 1,
                                            cp = 0.01))
```

```r
print(tree_model)
```

```
## n= 10
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 10 3 no (0.7000000 0.3000000)
##    2) Married_Status=married 4 0 no (1.0000000 0.0000000) *
##    3) Married_Status=divorcee,single 6 3 no (0.5000000 0.5000000)
##      6) Home_Owner=yes 2 0 no (1.0000000 0.0000000) *
##      7) Home_Owner=no 4 1 yes (0.2500000 0.7500000)
##       14) Annual_Income< 77500 1 0 no (1.0000000 0.0000000) *
##       15) Annual_Income>=77500 3 0 yes (0.0000000 1.0000000) *
```

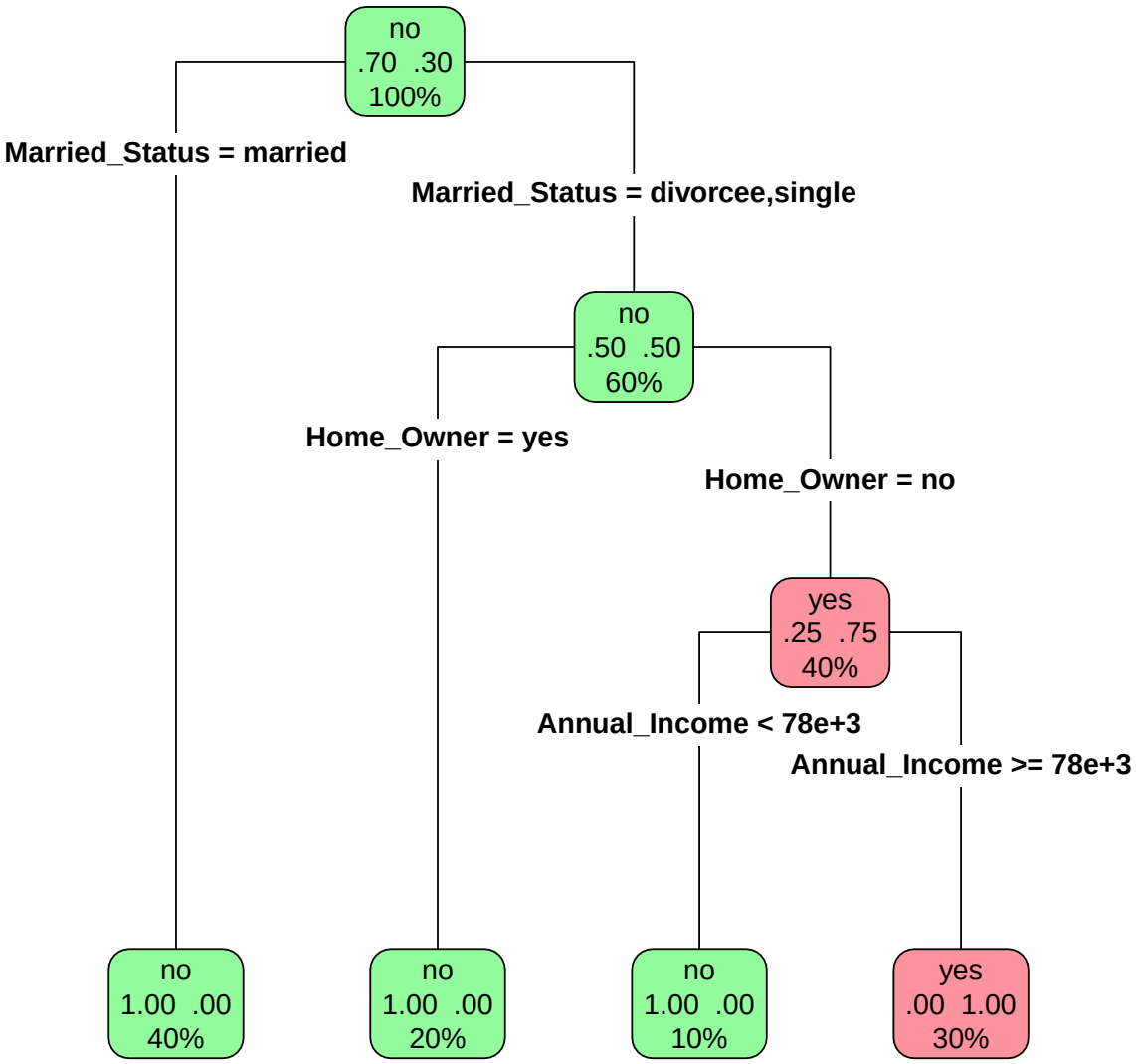
Visualizing the Decision Tree Model

```
rpart.plot(tree_model,
           type = 5,
           extra = 0,
           box.palette = c("#93fd9e", "#fd939d"))
```



The following diagram displays labels at all nodes, giving a comprehensive idea how the tree was made.

```
rpart.plot(tree_model,
           type = 4,
           extra = 104,
           clip.right.labs = FALSE,
           box.palette = c("#93fd9e", "#fd939d"))
```



Prediction on New Example

```
new_example <- data.frame(Home_Owner = "yes",
                          Married_Status = "divorcee",
                          Annual_Income = 120000)
```

```
predict(tree_model, new_example)

##   no yes
## 1  1   0
```

Our decision tree model predicts that a divorcee home owner with annual income 120000 dollars will not default in his loan.