

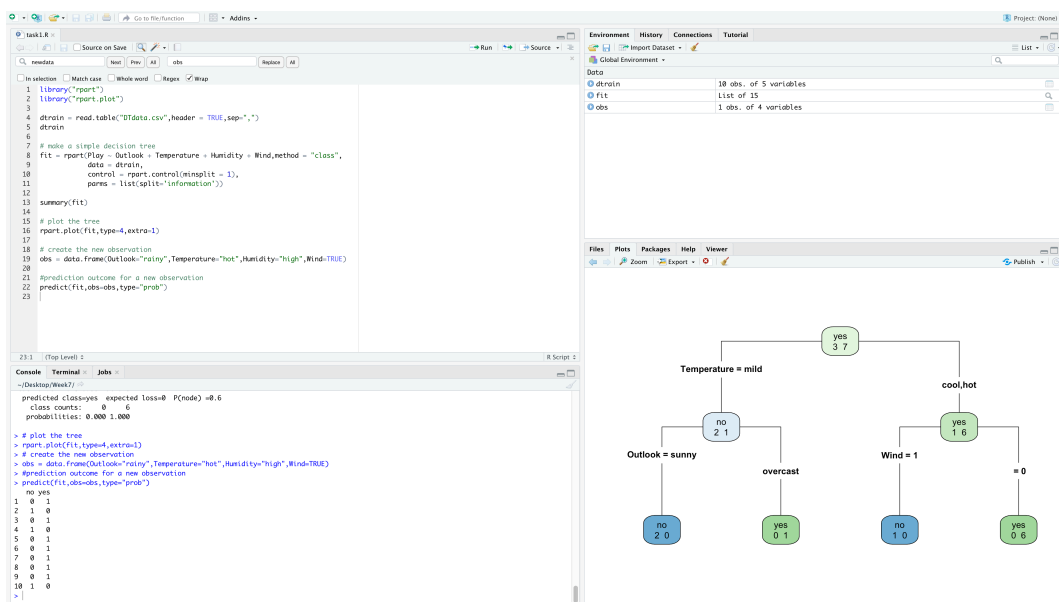
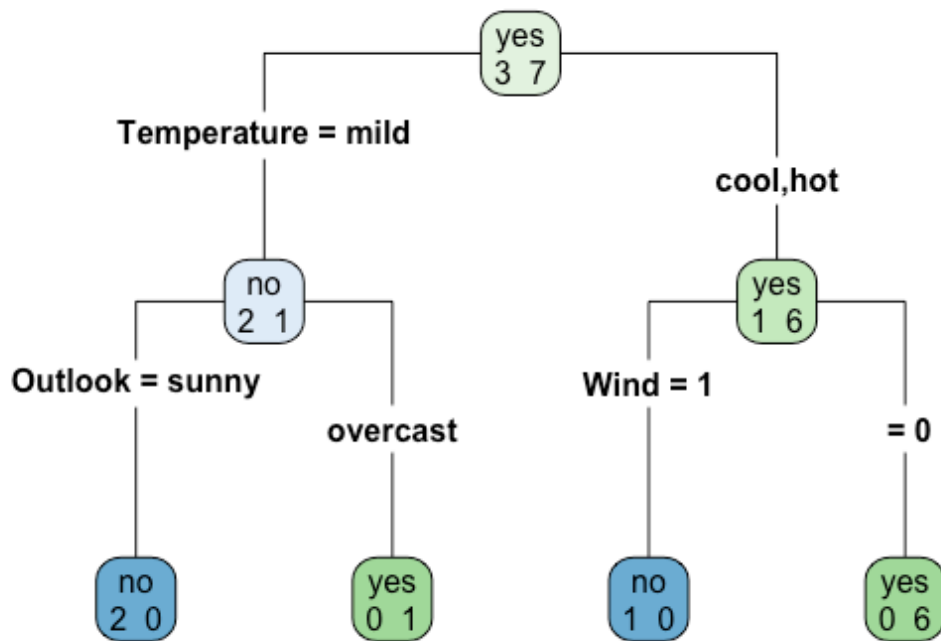
CSCI946 Assignment

Yao Xiao
SID 2019180015

November 3, 2020

1 Task One

```
1 library("rpart")
2 library("rpart.plot")
3
4 dtrain = read.table("DTdata.csv",header = TRUE,sep=",")
5 dtrain
6
7 # make a simple decision tree
8 fit = rpart(Play ~ Outlook + Temperature + Humidity + Wind,method =
    "class",
9             data = dtrain,
10            control = rpart.control(minsplit = 1),
11            parms = list(split='information'))
12
13 summary(fit)
14
15 # plot the tree
16 rpart.plot(fit,type=4,extra=1)
17
18 # create the new observation
19 obs = data.frame(Outlook="rainy",Temperature="hot",Humidity="high",
    Wind=TRUE)
20
21 #prediction outcome for a new observation
22 predict(fit,obs=obs,type="prob")
```



2 Task Two

```

1 library("e1071")
2
3 df = read.table("sample1.csv", header = TRUE, sep = ",")
4 df
5
6 traindata = as.data.frame(df[1:14,])
7 traindata
8 testdata = as.data.frame(df[15,])
9 testdata
10
11 # build model
12 model = naiveBayes(Enrolls ~ Age+Income+JobSatisfaction+Desire ,
13                   traindata, laplace = 0.01)
14 model
15 # make the prediction
16 results = predict(model, newdata = testdata, type = "raw")
17 results

```

The screenshot shows the RStudio IDE with the following components:

- Source Editor:** Contains the R code from Task Two.
- Console:** Displays the output of the code. It shows the structure of the data frames and the results of the naiveBayes model.


```

Income
  No 0.4633000 0.2820000 0.4633000
  Yes 0.2233333 0.3344444 0.4455556

JobSatisfaction
  No 0.8020000 0.2820000
  Yes 0.3344444 0.6677778

Desire
  Excellent 0.4633000
  Fair 0.3344444
  Poor 0.2022222

```
- Environment:** Shows the objects created during the execution: df, traindata, testdata, model, results, and traindata.
- Files:** Shows the project files: sample1.csv, sample1.csv.gz, and sample1.csv.gz.gz.