

## Practical 9

-Shivam Tawari A-58

Aim: Write a program in R for implementing a Decision Tree on given dataset.

### Theory:

Decision tree is a graph to represent choices and their results in form of a tree. The nodes in the graph represent an event or choice and the edges of the graph represent the decision rules or conditions. It is mostly used in Machine learning and Data mining applications using R.

Examples of use of decision trees is - predicting an email as spam or not spam, predicting if a tumor is cancerous or not, predicting a loan as a good or bad credit risk based on the factors in each of these. Generally, a model is created with observed data also called training data. Then a set of validation data is used to verify and improve the model. R has packages which are used to create and visualize decision trees. For new set of predictors

variable, we use this model to arrive at a decision on the category of the data.

- The R package "party" is used to create decision trees.

Code:

```
install.packages("party")  
library("party")  
print(head(readingSkills))  
input.dat <- readingSkills[c(1:105),]  
png(file = "decision_tree.png")  
output.tree <- ctree(  
  nativeSpeaker ~ age + shoeSize + score,  
  data = input.dat)  
plot(output.tree)  
dev.off()
```

Conclusion: Hence, we have successfully implemented a decision tree in R for the given dataset.

## Code:

```
1 # Shivam Tawari (A-58)
2 install.packages("party")
3 library("party")
4 print(head(readingSkills))
5 input.dat <- readingSkills[c(1:105),]
6 png(file = "decision_tree.png")
7 output.tree <- ctree(
8   nativeSpeaker ~ age + shoeSize + score,
9   data = input.dat)
10 plot(output.tree)
11 dev.off()
```

## Output:

```
Console Terminal x Jobs x
/cloud/project/
Loading required package: sandwich
> print(head(readingSkills))
  nativeSpeaker age shoeSize  score
1          yes   5  24.83189 32.29385
2          yes   6  25.95238 36.63105
3          no  11  30.42170 49.60593
4          yes   7  28.66450 40.28456
5          yes  11  31.88207 55.46085
6          yes  10  30.07843 52.83124
> input.dat <- readingSkills[c(1:105),]
> png(file = "decision_tree.png")
> output.tree <- ctree(
+   nativeSpeaker ~ age + shoeSize + score,
+   data = input.dat)
> plot(output.tree)
> dev.off()
null device
      1
```

