# Lab 3

# Decision Tree in R

### 1. Install R Package

Use the below command in R console to install the package. You also have to install the dependent packages if any.

```
> install.packages("party")
```

The package "party" has the function **ctree()** which is used to create and analyze decision tree.

### 2. Tree Syntax

The basic syntax for creating a decision tree in R is –

```
> ctree(formula, data)
```

### 3. Input Data

We will use the R in-built data set named **readingSkills** to create a decision tree. It describes the score of someone's readingSkills if we know the variables "age","shoesize","score" and whether the person is a native speaker or not.

Let's load the party package. It will automatically load other dependent packages, and then print some records from data set readingSkills.

Here is the sample data.

```
> library(party)
> 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
> |
```

### 4. Create a Decision Tree

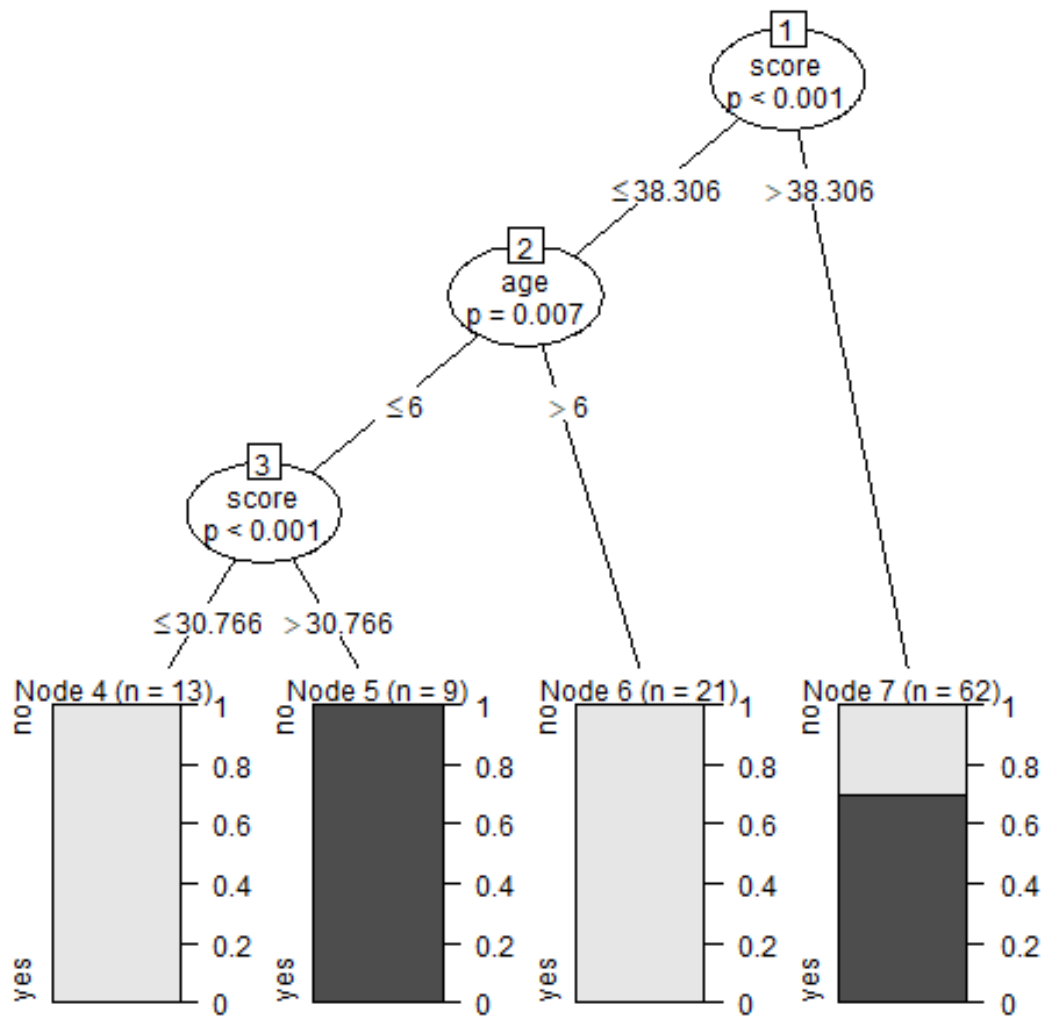We will use the **ctree()** function to create the decision tree and see its graph.

```
# Create the input data frame
input_data <- readingSkills[c(1:105),]

# Give a name to the chart file
png(file = "decision_tree.png")

# Now, create the tree
  output_tree <- ctree(
  nativeSpeaker ~ age + shoeSize + score,
  data = input_data)

# Now, plot the tree
plot(output_tree)

# Finally, save the file
dev.off()
```

## 5. Conclusions

What conclusions can we induce?

- From the decision tree shown above we can conclude that anyone whose readingSkills score is less than 38.3 and age is more than 6 is not a native Speaker.