

MODUL 13

POHON KEPUTUSAN 1 (SATU VARIABEL BEBAS)



CAPAIAN PEMBELAJARAN

1. Praktikan mampu mengimplementasikan penggunaan software R untuk analisis pohon keputusan dengan 1 variabel bebas



KEBUTUHAN ALAT/BAHAN/SOFTWARE

1. Komputer
2. Software R



DASAR TEORI

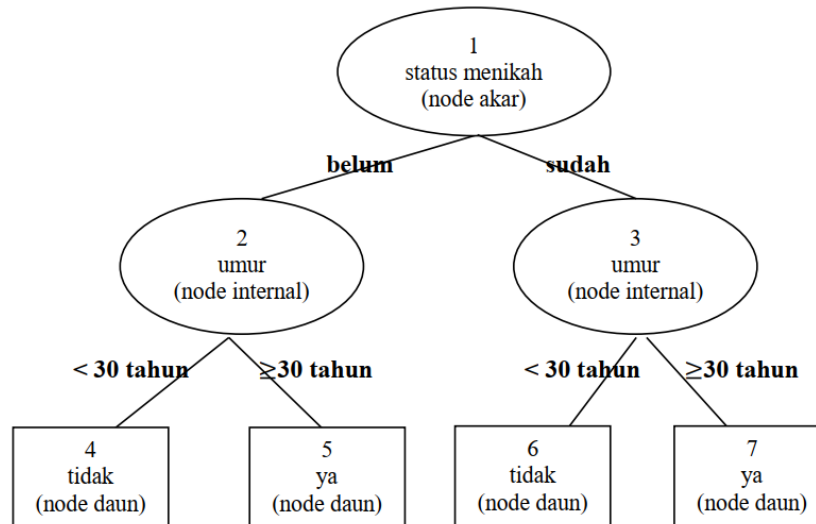
A. PENDAHULUAN

Berikut diberikan data mengenai catatan kepemilikan mobil

Nama	Umur	Status Menikah	Kepemilikan
A	25	sudah	ya
B	26	belum	tidak
C	28	belum	tidak
D	19	sudah	tidak
E	28	sudah	ya
F	40	sudah	ya
G	35	sudah	ya
H	32	belum	ya
I	33	sudah	tidak
J	55	sudah	ya

Berdasarkan data pada tabel di atas, diketahui responden bernama A, berusia 25 tahun, dengan status sudah menikah, memiliki mobil. Responden bernama H, berusia 32 tahun, dengan status belum menikah, memiliki mobil, dan seterusnya.

Berdasarkan data pada tabel tersebut, dibentuk pohon keputusan (decision tree) sebagai berikut.



Dalam R, pembuatan pohon keputusan dapat menggunakan fungsi **rpart** (pada library **rpart**) atau fungsi **tree** (pada library **tree**).



PRAKTIK

Praktik 1 (Library “rpart” dan “rpart.plot”)

Andaikan diberikan data pada Tabel berikut:

Tabel 1

Y	X ₁
A	3
A	1
A	5
B	9
B	12
B	7

Input

Inputkan data Tabel 1 pada tersebut kemudian aktifkan library **rpart** dan **rpart.plot** untuk membuat pohon keputusan.

```
D:\MATERI KULIAH\MODUL PRAKTIKUM STATISTIKA TERAPAN\PROGRAM\pohon1.R - R Editor
#menginputkan data
y <- c("A","A","A","B","B","B")
x1 <- c(3,1,5,9,12,7)
datapohon1 <- data.frame(y,x1)
datapohon1

#library rpart
library(rpart)
treel <- rpart(y~x1, datapohon1, minsplit=1)
summary(treel)
print(treel)

#library rpart.plot
prp(treel, faclen = 0, cex = 0.8, extra = 1)
treeprediction <- predict(treel, datapohon1, type = "class")
treeprediction
```

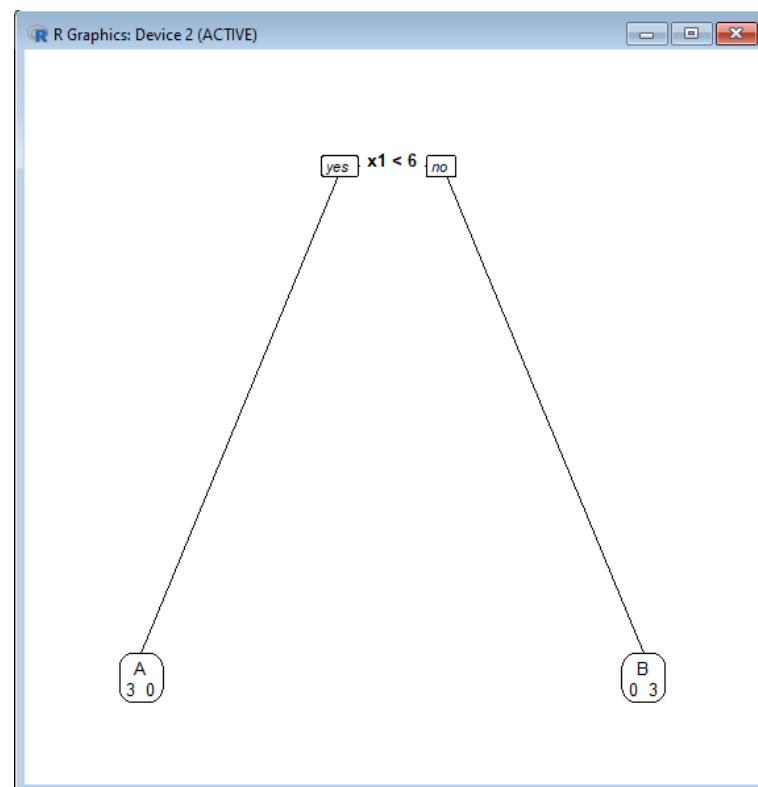
Gambar 1

Output

```
> print(treel)
n= 6

node), split, n, loss, yval, (yprob)
      * denotes terminal node

1) root 6 3 A (0.5000000 0.5000000)
  2) x1< 6 3 0 A (1.0000000 0.0000000) *
  3) x1>=6 3 0 B (0.0000000 1.0000000) *
```



Gambar 2

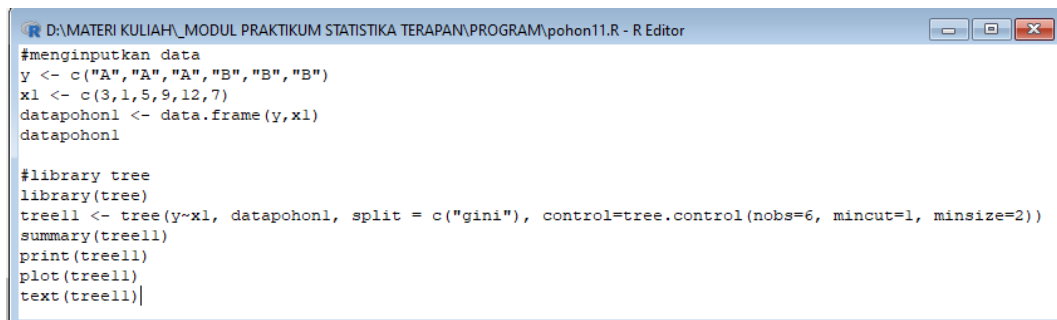
Analisis

Dari output di atas diperoleh **titik pemecahan optimal** terletak pada $x_1 = 6$, sehingga pada fungsi **rpart** digunakan pengelompokan $x_1 < 6$ dan $x_1 \geq 6$.

Praktik 2 (Library “tree”)

Input

Inputkan kembali data pada Tabel 1, kemudian aktifkan library **tree** untuk membuat pohon keputusan.



```
D:\MATERI KULIAH\MODUL PRAKTIKUM STATISTIKA TERAPAN\PROGRAM\pohon11.R - R Editor
#menginputkan data
y <- c("A","A","A","B","B","B")
x1 <- c(3,1,5,9,12,7)
datapohon1 <- data.frame(y,x1)
datapohon1

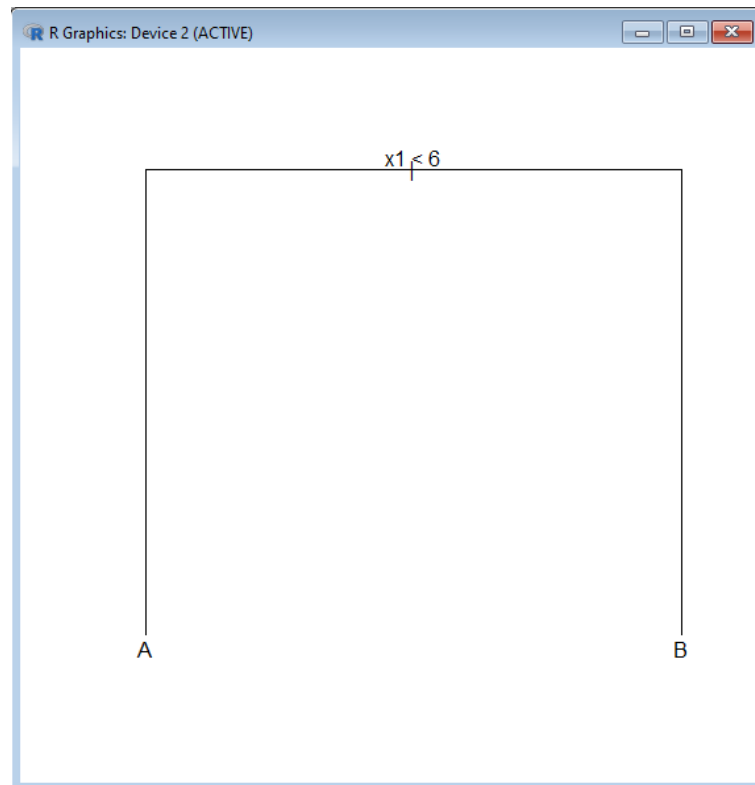
#library tree
library(tree)
treell <- tree(y~x1, datapohon1, split = c("gini"), control=tree.control(nobs=6, mincut=1, minsize=2))
summary(treell)
print(treell)
plot(treell)
text(treell)
```

Gambar 3

Output

```
> print(treell)
node), split, n, deviance, yval, (yprob)
      * denotes terminal node

1) root 6 8.318 A ( 0.5 0.5 )
  2) x1 < 6 3 0.000 A ( 1.0 0.0 ) *
  3) x1 > 6 3 0.000 B ( 0.0 1.0 ) *
```



Gambar 4

Analisis

Dari output di atas diperoleh **titik pemecahan optimal** terletak pada $x_1 = 6$, sehingga pada fungsi **tree** digunakan pengelompokan $x_1 < 6$ dan $x_1 > 6$. (Bandingkan dengan fungsi **rpart**)

Praktik 3 (Library “rpart” dan “rpart.plot”)

Andaikan diberikan data pada Tabel berikut:

Tabel 2

Y	X ₁
A	3
A	1
A	5
B	9
A	12
B	7

Input

Inputkan data Tabel 2 pada tersebut kemudian aktifkan library **rpart** dan **rpart.plot** untuk membuat pohon keputusan.

```

D:\MATERI KULIAH\MODUL PRAKTIKUM STATISTIKA TERAPAN\PROGRAM\pohon2.R - R Editor
#menginputkan data
y <- c("A","A","A","B","A","B")
x1 <- c(3,1,5,9,12,7)
datapohon2 <- data.frame(y,x1)
datapohon2

#library rpart
library(rpart)
tree2 <- rpart(y~x1, datapohon2, minsplit=1)
summary(tree2)
print(tree2)

#library rpart.plot
prp(tree2, facilen = 0, cex = 0.8, extra = 1)
treeprediction <- predict(tree2, datapohon2, type = "class")
treeprediction

```

Gambar 5

Output

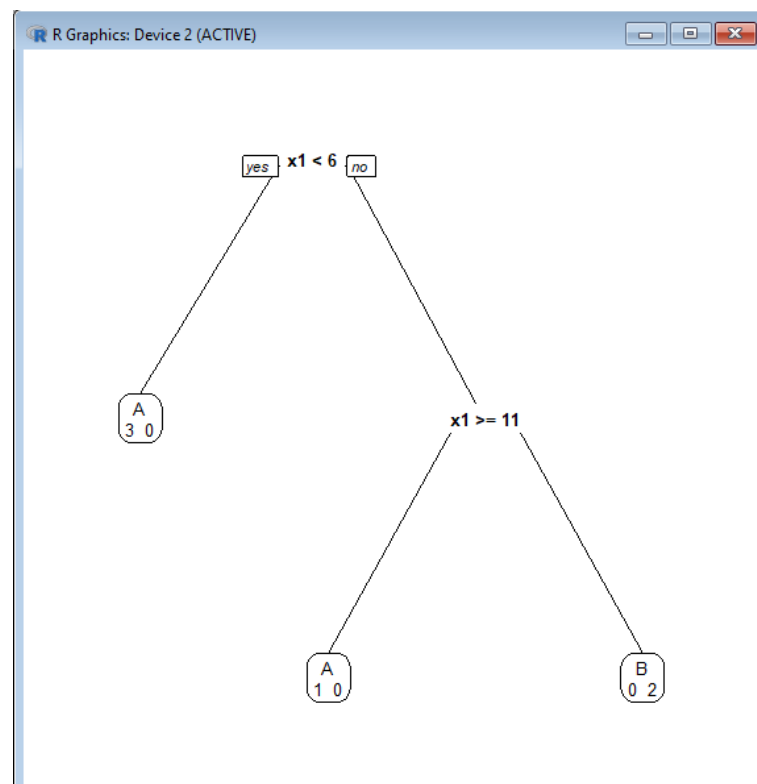
```

> print(tree2)
n= 6

node), split, n, loss, yval, (yprob)
      * denotes terminal node

1) root 6 2 A (0.6666667 0.3333333)
 2) x1< 6 3 0 A (1.0000000 0.0000000) *
 3) x1>=6 3 1 B (0.3333333 0.6666667)
   6) x1>=10.5 1 0 A (1.0000000 0.0000000) *
   7) x1< 10.5 2 0 B (0.0000000 1.0000000) *
> |

```



Gambar 6

Analisis

Dari output di atas diperoleh **titik pemecahan optimal** dengan fungsi **rpart** terletak pada $x_1 = 6$ pada level 1 dan $x_1 \geq 11$ pada level 2.

Praktik 4 (Library “tree”)

Input

Inputkan kembali data pada Tabel 2, kemudian aktifkan library **tree** untuk membuat pohon keputusan.



```
D:\MATERI KULIAH\MODUL PRAKTIKUM STATISTIKA TERAPAN\PROGRAM\pohon21.R - R Editor
#menginputkan data
y <- c("A","A","A","B","A","B")
x1 <- c(3,1,5,9,12,7)
datapohon2 <- data.frame(y,x1)
datapohon2

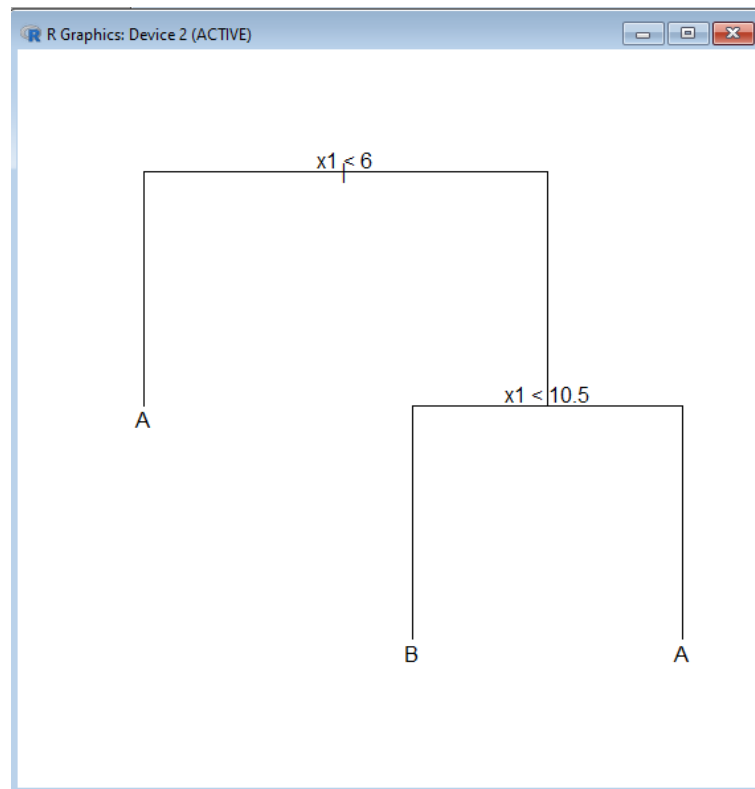
#library tree
library(tree)
tree21 <- tree(y~x1, datapohon2, split = c("gini"), control=tree.control(nobs=6, mincut=1, minsize=2))
summary(tree21)
print(tree21)
plot(tree21)
text(tree21)
```

Gambar 7

Output

```
> print(tree21)
node), split, n, deviance, yval, (yprob)
      * denotes terminal node

1) root 6 7.638 A ( 0.6667 0.3333 )
 2) x1 < 6 3 0.000 A ( 1.0000 0.0000 ) *
 3) x1 > 6 3 3.819 B ( 0.3333 0.6667 )
    6) x1 < 10.5 2 0.000 B ( 0.0000 1.0000 ) *
    7) x1 > 10.5 1 0.000 A ( 1.0000 0.0000 ) *
```



Gambar 8

Analisis

Dari output di atas diperoleh **titik pemecahan optimal** dengan fungsi **tree** terletak pada $x_1 = 6$ pada level 1 dan $x_1 < 10.5$ pada level 2.



LATIHAN

Y	X1
A	3
A	1
A	2
A	3
B	9
B	8
B	10
B	14
A	13
A	12

Buatlah pohon keputusan dari data tersebut menggunakan fungsi **rpart** dan **tree**, kemudian tentukan titik pemecah optimal dan pengelompokannya!



TUGAS

Y	X1
A	25
A	27
A	4
A	1
B	20
B	24
B	23
B	21
A	3
A	13

Buatlah pohon keputusan dari data tersebut menggunakan fungsi **rpart** dan **tree**, kemudian tentukan titik pemecah optimal dan pengelompokannya!



REFERENSI

- [1] Gio, P.U., Effendie, A.R. 2017. Belajar Bahasan Pemrograman R (Dilengkapi Cara Membuat Aplikasi Olah Data Sederhana dengan R Shiny). Medan: USU Press.