

Prezentacja i wizualizacja danych

Organizacyjnie

Prowadzący:

dr Mariusz Rafało

mrafalo@sggw.edu.pl

<http://mariuszrafalo.pl> (hasło:WIZ)

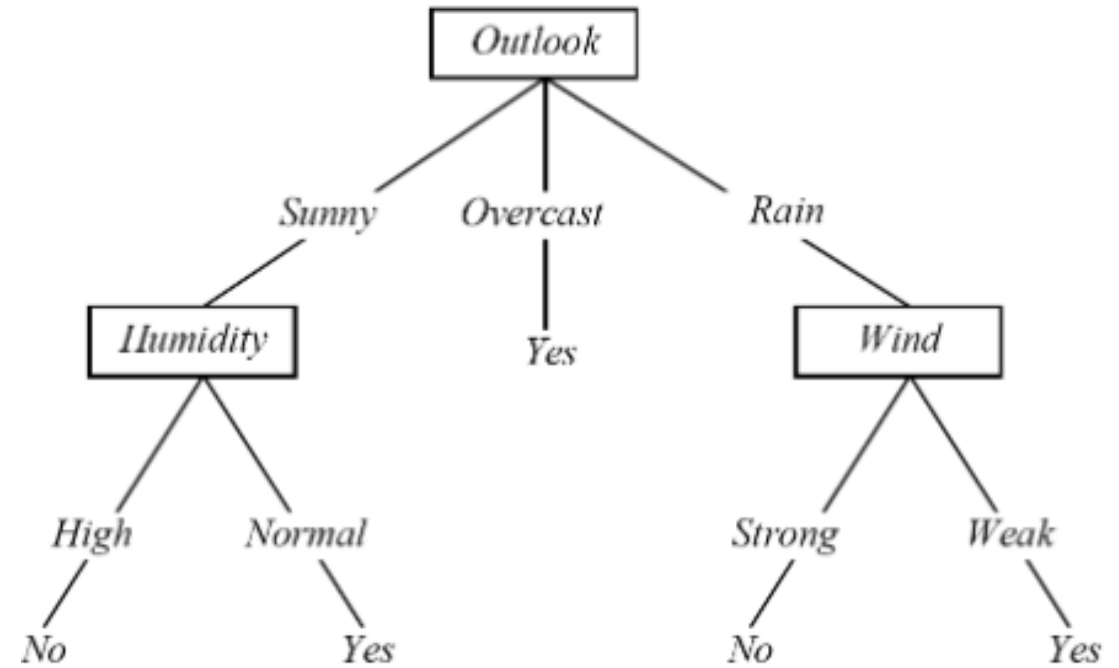
Drzewa decyzyjne

Drzewa decyzyjne

- Drzewo decyzyjne to metoda analitycznego i graficznego wspomaganie decyzji
- W modelach predykcyjnych technika drzew decyzyjnych służy identyfikacji reguł (gałęzi), prowadzących do zajścia określonego zdarzenia

Drzewo decyzyjne – przykład

| Outlook | Temperature | Humidity | Windy | PlayTennis |
|----------|-------------|----------|-------|------------|
| Sunny | Hot | High | False | No |
| Sunny | Hot | High | True | No |
| Overcast | Hot | High | False | Yes |
| Rainy | Mild | High | False | Yes |
| Rainy | Cool | Normal | False | Yes |
| Rainy | Cool | Normal | True | No |
| Overcast | Cool | Normal | True | Yes |
| Sunny | Mild | High | False | No |
| Sunny | Cool | Normal | False | Yes |
| Rainy | Mild | Normal | False | Yes |
| Sunny | Mild | Normal | True | Yes |
| Overcast | Mild | High | True | Yes |
| Overcast | Hot | Normal | False | Yes |
| Rainy | Mild | High | True | No |



Drzewa decyzyjne - R

```
library(rpart)

model.dt <- rpart(Churn ~ AccountLength + DayMins + IntlPlan + VMailPlan +
IntlCalls + CustServCalls + EveCalls + NightCalls + VMailMessage, data=dane.train)

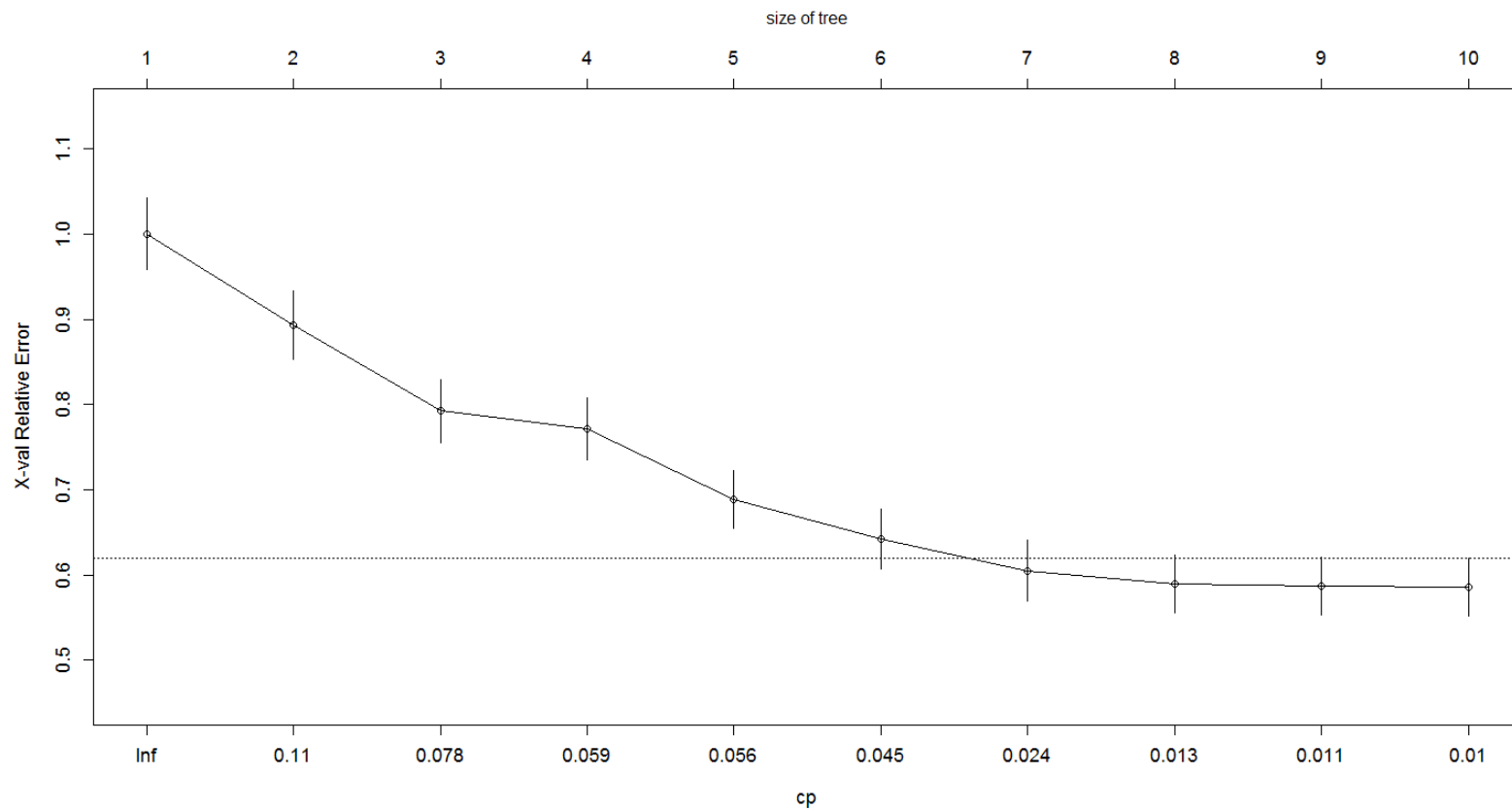
printcp(model.dt)
```

Drzewa decyzyjne – R: diagnostyka

| | CP | nsplit | rel error | xerror | xstd |
|----|----------|--------|-----------|---------|----------|
| 1 | 0.121965 | 0 | 1.00000 | 1.00177 | 0.042246 |
| 2 | 0.101204 | 1 | 0.87804 | 0.91312 | 0.041206 |
| 3 | 0.059567 | 2 | 0.77683 | 0.79628 | 0.037013 |
| 4 | 0.059376 | 3 | 0.71726 | 0.77071 | 0.036240 |
| 5 | 0.052521 | 4 | 0.65789 | 0.72536 | 0.035802 |
| 6 | 0.037994 | 5 | 0.60537 | 0.65325 | 0.034882 |
| 7 | 0.015600 | 6 | 0.56737 | 0.61514 | 0.035472 |
| 8 | 0.010735 | 7 | 0.55177 | 0.59625 | 0.033674 |
| 9 | 0.010641 | 8 | 0.54104 | 0.59375 | 0.033484 |
| 10 | 0.010000 | 9 | 0.53040 | 0.59180 | 0.033307 |

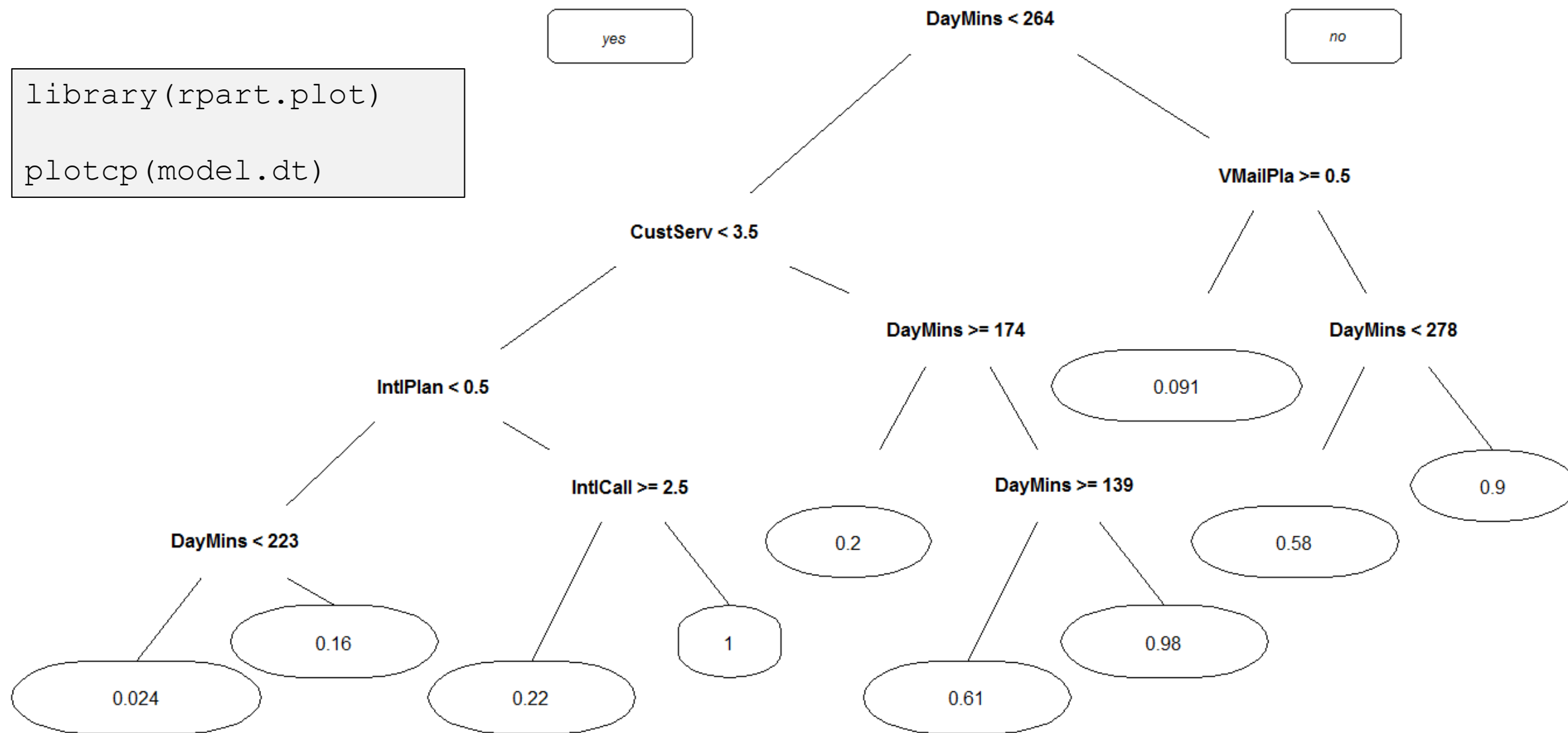
Drzewa decyzyjne - R

```
plotcp(model.dt, cex.lab=1.1, cex.axis=1.1)
```



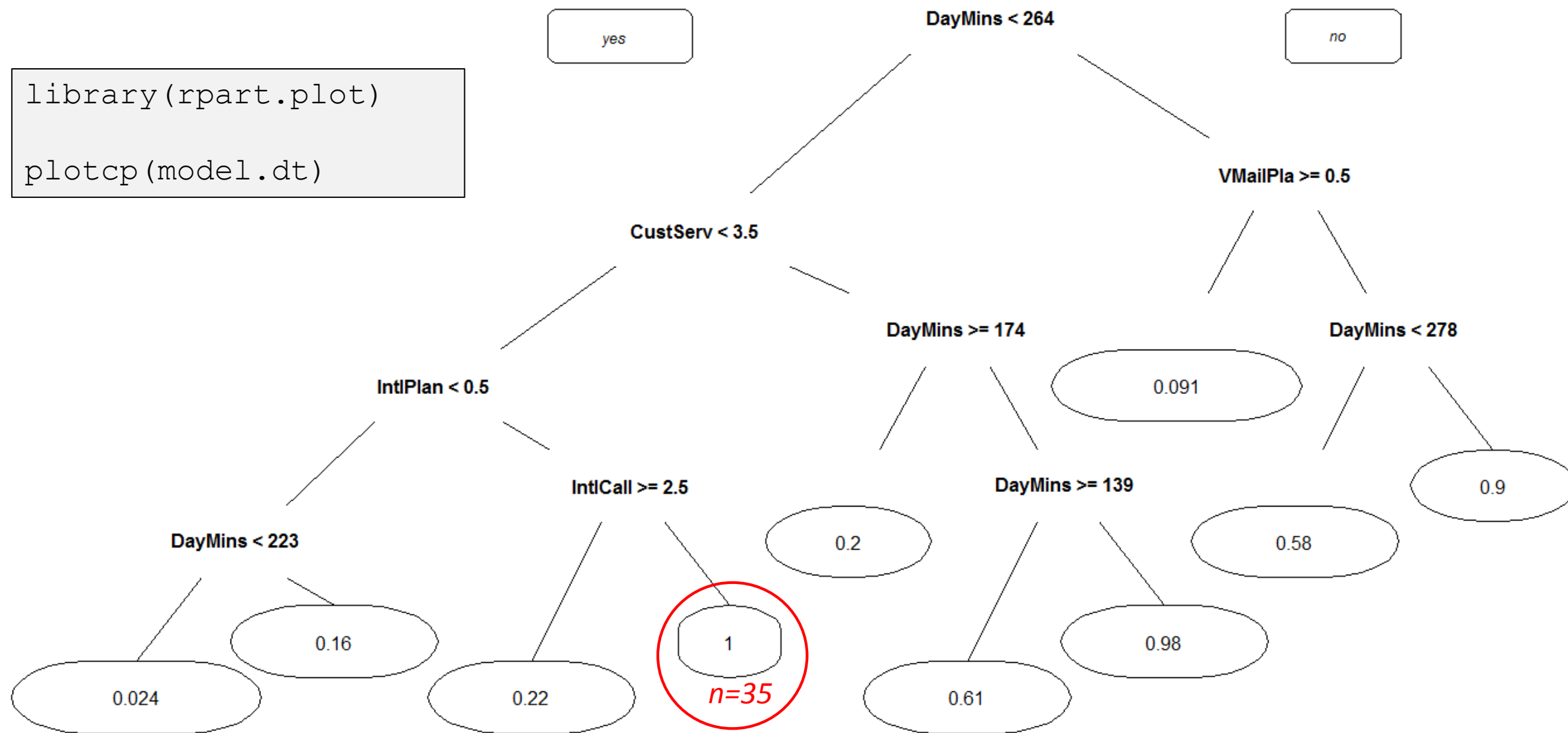
Drzewa decyzyjne - R

```
library(rpart.plot)
plotcp(model.dt)
```



Drzewa decyzyjne - R

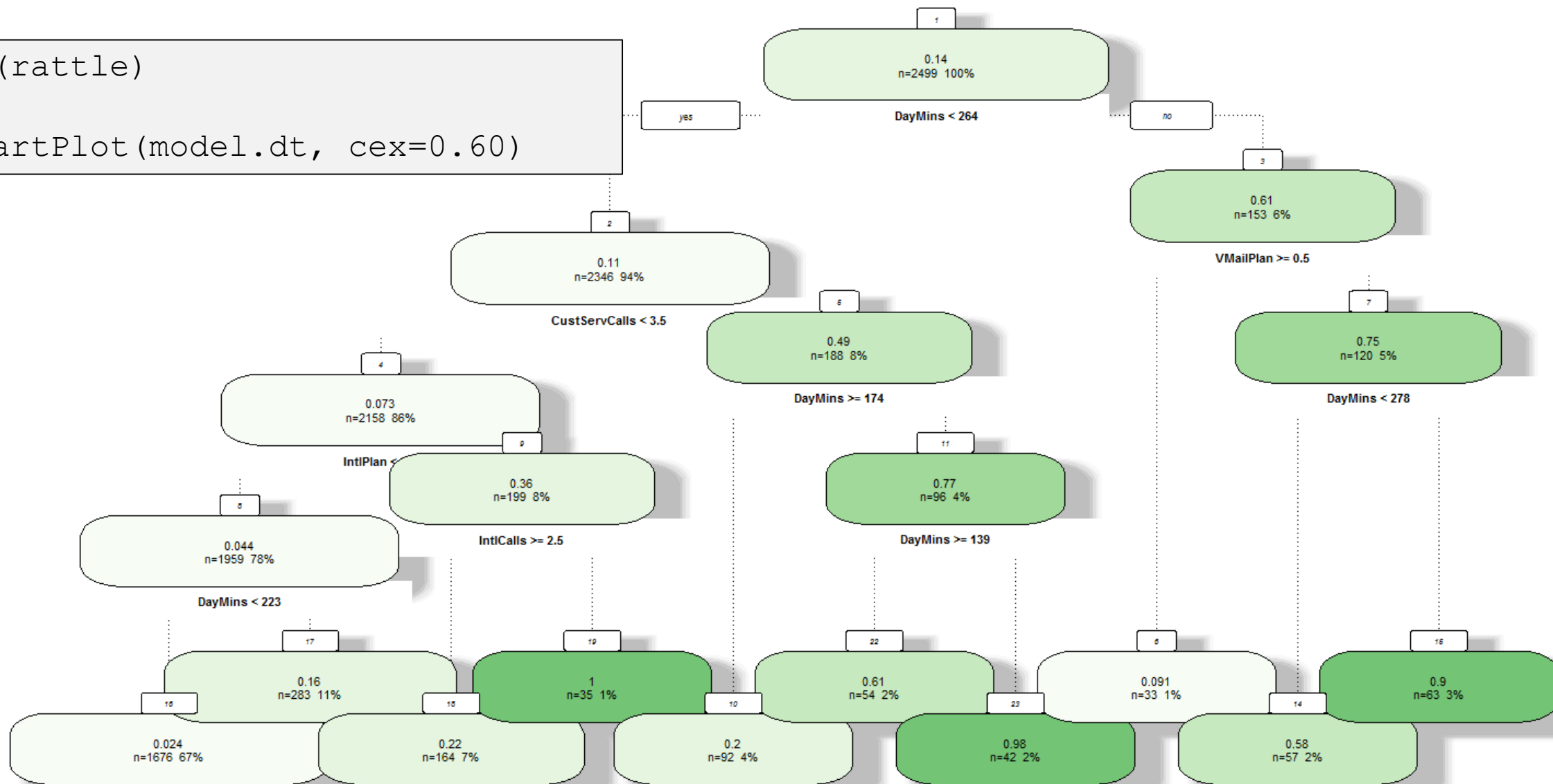
```
library(rpart.plot)
plotcp(model.dt)
```



Drzewa decyzyjne - R

```
library(rattle)
```

```
fancyRpartPlot(model.dt, cex=0.60)
```



Dziękuję za uwagę