

No :

Date :

Nama : M. Hafiz Al Ihsan

NIM : 221810407

No. Absen : 16

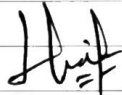
Kelas : 3SI2

Dosen : Ibnu Santoso, SST., MT.

Mata Kuliah : DMKM

Hari / Tanggal Ujian : 27 Oktober 2020 / Selasa

"Saya menyatakan bahwa ujian ini saya kerjakan dengan jujur sesuai kemampuan sendiri dan tidak mengutip sebagian atau seluruh pekerjaan orang lain. Jika suatu saat ditemukan saya melanggar ketentuan ujian, saya siap menerima konsekuensi yang berlaku."



M. Hafiz Al Ihsan



DMKN

M. Hafiz Al Ihsan

October 27, 2020

Library

```
library(party)
library(psych)
library(caret)
library(dplyr)
```

Import Data

```
library(readr)
kredit <- read_csv("C:/Users/Hafiz/OneDrive/Documents/PJJ/Tingkat 3/UTS/DMKN/
crx.csv")
```

```
## Parsed with column specification:
```

```
## cols(
##   A1 = col_character(),
##   A2 = col_character(),
##   A3 = col_double(),
##   A4 = col_character(),
##   A5 = col_character(),
##   A6 = col_character(),
##   A7 = col_character(),
##   A8 = col_double(),
##   A9 = col_logical(),
##   A10 = col_logical(),
##   A11 = col_character(),
##   A12 = col_logical(),
##   A13 = col_character(),
##   A14 = col_character(),
##   A15 = col_double(),
##   A16 = col_character()
## )
```

```
head(kredit)
```

```
## # A tibble: 6 x 16
```

```
##   A1      A2      A3 A4      A5      A6      A7      A8 A9      A10      A11      A12
##   <chr> <chr> <dbl> <chr> <chr> <chr> <chr> <dbl> <lgl> <lgl> <chr> <lgl>
## 1 b      30.83  0      u      g      w      v      1.25 TRUE  TRUE  01     FALSE
## 2 a      58.67  4.46 u      g      q      h      3.04 TRUE  TRUE  06     FALSE
```

```
g
## 3 a      24.50  0.5  u      g      q      h      1.5  TRUE  FALSE 0      FALSE
g
## 4 b      27.83  1.54 u      g      w      v      3.75 TRUE  TRUE  05     TRUE
g
## 5 b      20.17  5.62 u      g      w      v      1.71 TRUE  FALSE 0      FALSE
s
## 6 b      32.08  4    u      g      m      v      2.5  TRUE  FALSE 0      TRUE
g
## # ... with 3 more variables: A14 <chr>, A15 <dbl>, A16 <chr>
```

Missing Value

```
#Menghitung Jumlah Missing
sum(is.na(kredit))
```

```
## [1] 0
```

Missing value tidak ada

Kalsifikasi Decision Tree

1. Konversi Data

Ubah tipe variabel menjadi tipe faktor

```
kredit <- kredit %>% mutate_if(is.character,as.factor)
str(kredit)

## tibble [690 x 16] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ A1 : Factor w/ 3 levels "?","a","b": 3 2 2 3 3 3 3 2 3 3 ...
##  $ A2 : Factor w/ 350 levels "?","13.75","15.17",...: 158 330 91 127 45 170
## 181 76 312 257 ...
##  $ A3 : num [1:690] 0 4.46 0.5 1.54 5.62 ...
##  $ A4 : Factor w/ 4 levels "?","l","u","y": 3 3 3 3 3 3 3 3 4 4 ...
##  $ A5 : Factor w/ 4 levels "?","g","gg","p": 2 2 2 2 2 2 2 2 4 4 ...
##  $ A6 : Factor w/ 15 levels "?","aa","c","cc",...: 14 12 12 14 14 11 13 4 1
## 0 14 ...
##  $ A7 : Factor w/ 10 levels "?","bb","dd",...: 9 5 5 9 9 9 5 9 5 9 ...
##  $ A8 : num [1:690] 1.25 3.04 1.5 3.75 1.71 ...
##  $ A9 : logi [1:690] TRUE TRUE TRUE TRUE TRUE TRUE TRUE ...
##  $ A10: logi [1:690] TRUE TRUE FALSE TRUE FALSE FALSE ...
##  $ A11: Factor w/ 23 levels "0","01","02",...: 2 7 1 6 1 1 1 1 1 1 ...
##  $ A12: logi [1:690] FALSE FALSE FALSE TRUE FALSE TRUE ...
##  $ A13: Factor w/ 3 levels "g","p","s": 1 1 1 1 3 1 1 1 1 1 ...
##  $ A14: Factor w/ 171 levels "?","00000","00017",...: 70 13 98 33 39 117 56
## 25 64 17 ...
##  $ A15: num [1:690] 0 560 824 3 0 ...
##  $ A16: Factor w/ 2 levels "-","+": 2 2 2 2 2 2 2 2 2 2 ...
##  - attr(*, "spec")=
```

```
## .. cols(
## ..   A1 = col_character(),
## ..   A2 = col_character(),
## ..   A3 = col_double(),
## ..   A4 = col_character(),
## ..   A5 = col_character(),
## ..   A6 = col_character(),
## ..   A7 = col_character(),
## ..   A8 = col_double(),
## ..   A9 = col_logical(),
## ..   A10 = col_logical(),
## ..   A11 = col_character(),
## ..   A12 = col_logical(),
## ..   A13 = col_character(),
## ..   A14 = col_character(),
## ..   A15 = col_double(),
## ..   A16 = col_character()
## .. )
```

2. Split Data

Memecah data menjadi data training (80% dari data awal) dan data test (20% dari data awal)

```
set.seed(1234)
sampel <- sample(2,nrow(kredit),replace = T, prob = c(0.8,0.2))
trainingdat <- kredit[sampel==1, ]
testingdat <- kredit[sampel==2, ]
print(paste("Jumlah train data :", nrow(trainingdat)))

## [1] "Jumlah train data : 544"

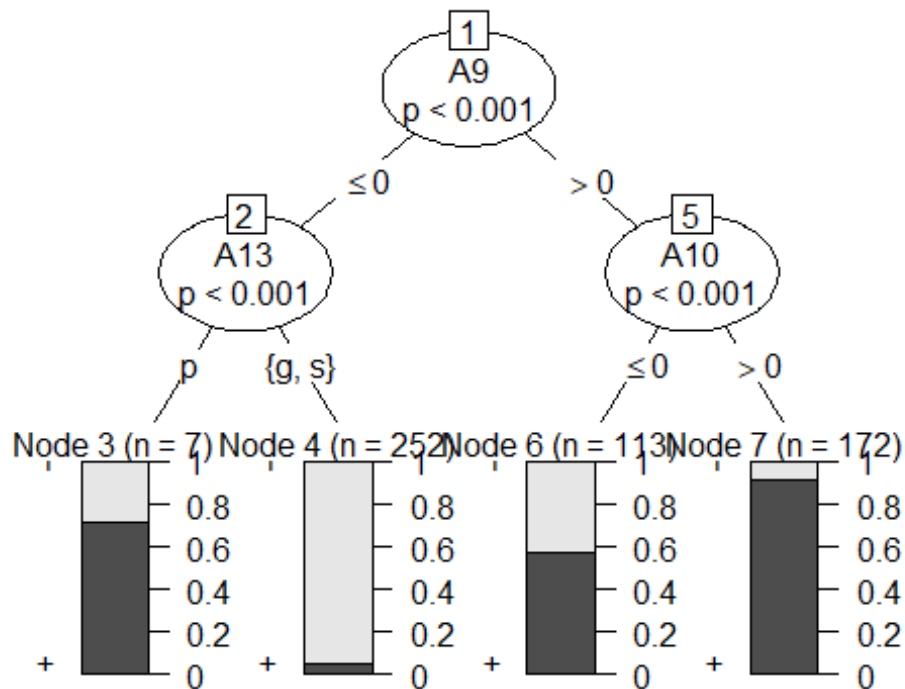
print(paste("Jumlah test data :", nrow(testingdat)))

## [1] "Jumlah test data : 146"
```

3. Membuat Model

Kita membuat model berdasarkan variabel bertipe nominal saja

```
decision_tree <- ctree(A16~A1+A4+A5+A6+A7+A9+A10+A12+A13, data=trainingdat)
plot(decision_tree)
```



4. Model Evaluation

```
prediksi <- predict(decision_tree, testingdat)
confusionMatrix(table(prediksi, testingdat$A16))
```

```
## Confusion Matrix and Statistics
##
##
## prediksi  -  +
##          - 65  4
##          + 15 62
##
##              Accuracy : 0.8699
##              95% CI : (0.8043, 0.9198)
##      No Information Rate : 0.5479
##      P-Value [Acc > NIR] : < 2e-16
##
##              Kappa : 0.7411
##
##  Mcnemar's Test P-Value : 0.02178
##
##              Sensitivity : 0.8125
##              Specificity : 0.9394
##      Pos Pred Value : 0.9420
##      Neg Pred Value : 0.8052
##              Prevalence : 0.5479
##      Detection Rate : 0.4452
```

```
## Detection Prevalence : 0.4726
## Balanced Accuracy : 0.8759
##
## 'Positive' Class : -
##
```

5. Interpretasi Statistik Evaluasi

- Sebanyak 65 berhasil diprediksi oleh model untuk pemberian kredit - dan sebanyak 62 berhasil diprediksi oleh model untuk pemberian kredit +
- Pemberian kredit benar diprediksi oleh model dengan tingkat akurasi 0.8699
- Pemberian kredit yang bernilai - dibandingkan dengan keseluruhan data pemberian kredit yang sebenarnya bernilai - tepat diprediksi sebanyak 81.25%
- Pemberian kredit yang bernilai + dibandingkan dengan keseluruhan data pemberian kredit yang sebenarnya bernilai + tepat diprediksi sebanyak 93.94%
- F1-Score = 87.135%