

PROJECT LOAN CLASSIFICATION ANALYSIS FOR UMKM

BACKGROUND

Perusahaan Alembert merupakan perusahaan yang bergerak di bidang layanan pinjaman usaha bagi sektor UMKM. Karena adanya pandemik ini, perusahaan berusaha memberikan pelayanan berupa keringanan pinjaman bagi pelanggan yang disebut sebagai rekomendasi tindak lanjut. Pemberian rekomendasi tindak lanjut pada pelanggan ini didasari pada kriteria tertentu, dan perlu ditentukan faktor-faktor apa saja yang berpengaruh sehingga pelanggan mendapatkan treatment tertentu yang masuk dalam rekomendasi tindak lanjut program dari perusahaan.

Tujuan Project: Klasifikasi nasabah yang akan dimasukkan pada rekomendasi tindak lanjut. Pada kelas target rekomendasi tindak lanjut ini sendiri terdiri dari beberapa kelas seperti restrukturisasi dan angsuran biasa.

Model: Regresi multinomial.

Referensi model (Ref): Kelas pada rekomendasi tindak lanjut yang memiliki banyak pelanggan.

Data: Data yang digunakan terdiri dari 1000 baris.

READ EXTERNAL DATA

```
1 data = read.csv("https://storage.googleapis.com/dqlab-dataset/project.csv")
```

DATA INSPECTION

Setelah data berhasil di import, cobalah kamu untuk menginspeksi dataset dengan jalan

1. Melihat 6 baris pertama data tersebut, apa saja yang ditunjukkannya dan
2. Tampilkan tipe data dari setiap kolom.

```
1 # Enam baris teratas data
2 head(data)
3 # Tampilkan tipe data setiap kolomnya
4 str(data)
```

Output after input syntax above:

```
> # Enam baris teratas data
> head(data)
  X      NAMA_NASABAH NOMOR_KONTRAK      DOMISILI      KARAKTER
1 0 YOLI SEPINA NAINGGOLAN          0      MASIH TETAP      KOOPERATIF
2 1      ERWIN NASUTION          1      MASIH TETAP TIDAK      KOOPERATIF
3 2      HUSIN          2      MASIH TETAP TIDAK      KOOPERATIF
4 3      HARITSYAH          3 PINDAH PERMANEN      KOOPERATIF
5 4      HARIRI PANGGABEAN          4      MASIH TETAP TIDAK      KOOPERATIF
6 5      JHON PREDDY HUTABARAT          5      MASIH TETAP      KOOPERATIF
```

```

      PROFESI KONDISI_USAHA KONDISI_JAMINAN STATUS PRODUK      PYD
1      IBU RUMAH TANGGA      2      Baik      2      3 30000000
2      NELAYAN      3      Rusak      8      3 10000000
3      LAINNYA      3      Baik      8      3 60000000
4      PNS      1      Rusak      7      3 90000000
5 WIRSAUSAHA / PEDAGANG      1      Baik      8      3 150000000
6 WIRSAUSAHA / PEDAGANG      1      Baik      7      3 40000000
      TENOR      OSL KEWAJIBAN      KOLEKTIBILITAS COUNT_SURVEY
1      24 28750000 4896841 DALAM PENGAWASAN KHUSUS      1
2      12 2040693      0      MACET      1
3      24      0      0      MACET      1
4      18      0      0      MACET      1
5      18 19844807      0      MACET      1
6      36 27298726 2208516 DALAM PENGAWASAN KHUSUS      1
      REKOMENDASI_TINDAK_LANJUT
1      Angsuran Biasa
2      Penarikan
3      Penarikan
4      Angsuran Biasa
5      Penarikan
6      Restrukturisasi

> # Tampilkan tipe data setiap kolomnya
> str(data)
'data.frame':      1000 obs. of      17 variables:
 $ X      : int      0 1 2 3 4 5 6 7 8 9 ...
 $ NAMA_NASABAH      : chr      "YOLI SEPINA NAINGGOLAN" "ERWIN NASUTION" "HUSIN" "HARITSY
AH" ...
 $ NOMOR_KONTRAK      : int      0 1 2 3 4 5 6 7 8 9 ...
 $ DOMISILI      : chr      "MASIH TETAP" "MASIH TETAP" "MASIH TETAP" "PINDAH PERMANEN
" ...
 $ KARAKTER      : chr      "KOOPERATIF" "TIDAK KOOPERATIF" "TIDAK KOOPERATIF" "KOOPER
ATIF" ...
 $ PROFESI      : chr      "IBU RUMAH TANGGA" "NELAYAN" "LAINNYA" "PNS" ...
 $ KONDISI_USAHA      : int      2 3 3 1 1 1 3 2 2 3 ...
 $ KONDISI_JAMINAN      : chr      "Baik" "Rusak" "Baik" "Rusak" ...
 $ STATUS      : int      2 8 8 7 8 7 7 7 7 3 ...
 $ PRODUK      : int      3 3 3 3 3 3 3 3 3 3 ...
 $ PYD      : int      30000000 10000000 60000000 90000000 150000000 40000000 600
00000 7500000 45000000 50000000 ...
 $ TENOR      : int      24 12 24 18 18 36 36 4 24 36 ...
 $ OSL      : int      28750000 2040693 0 0 19844807 27298726 19999200 7500000 45
000000 6944100 ...
 $ KEWAJIBAN      : int      4896841 0 0 0 0 2208516 6946592 0 0 7730984 ...
 $ KOLEKTIBILITAS      : chr      "DALAM PENGAWASAN KHUSUS" "MACET" "MACET" "MACET" ...
 $ COUNT_SURVEY      : int      1 1 1 1 1 1 2 2 1 1 ...
 $ REKOMENDASI_TINDAK_LANJUT: chr      "Angsuran Biasa" "Penarikan" "Penarikan" "Angsuran Biasa"
...

```

DESCRIPTIVE STATISTICS

See the summary of data

```
1 summary(data)
```

Output after input syntax above:

```
> summary(data)
      X      NAMA_NASABAH      NOMOR_KONTRAK      DOMISILI
Min.   : 0.0      Length:1000      Min.   : 0.0      Length:1000
1st Qu.:249.8      Class :character      1st Qu.:249.8      Class :character
Median :499.5      Mode  :character      Median :499.5      Mode  :character
Mean   :499.5
3rd Qu.:749.2
Max.   :999.0

      KARAKTER      PROFESI      KONDISI_USAHA      KONDISI_JAMINAN
Length:1000      Length:1000      Min.   :1.000      Length:1000
Class :character      Class :character      1st Qu.:2.000      Class :character
Mode  :character      Mode  :character      Median :2.000      Mode  :character
                               Mean   :2.273
                               3rd Qu.:3.000
                               Max.   :3.000

      STATUS      PRODUK      PYD      TENOR
Min.   :2.000      Min.   : 3.00      Min.   : 500000      Min.   : 3.00
1st Qu.:3.000      1st Qu.: 3.00      1st Qu.: 8000000      1st Qu.:12.00
Median :7.000      Median : 3.00      Median : 15000000      Median :18.00
Mean   :5.379      Mean   :12.45      Mean   : 38537508      Mean   :20.75
3rd Qu.:7.000      3rd Qu.: 7.00      3rd Qu.: 50000000      3rd Qu.:24.00
Max.   :8.000      Max.   :77.00      Max.   :500000000      Max.   :48.00

      OSL      KEWAJIBAN      KOLEKTIBILITAS      COUNT_SURVEY
Min.   :      0      Min.   :      0      Length:1000      Min.   :1.000
1st Qu.: 3999950      1st Qu.: 687487      Class :character      1st Qu.:1.000
Median : 8687350      Median : 2008974      Mode  :character      Median :1.000
Mean   : 26562373      Mean   : 5663981
3rd Qu.: 32082900      3rd Qu.: 4823198
Max.   :440932336      Max.   :400900000
REKOMENDASI_TINDAK_LANJUT
Length:1000
Class :character
Mode  :character
```

DELETE COLUMN

Pada data, sebenarnya tidak memerlukan nama pelanggan untuk diberikan rekomendasi. Atau dengan kata lain penanda pelanggan untuk diberikan rekomendasi cukup dengan melihat no_kontrak pelanggan itu saja.

```
1 data_reduce = data[-c(1,2)]
2 colnames(data_reduce)
```

Output after input syntax above:

```
> data_reduce = data[-c(1,2)]  
  
> colnames(data_reduce)  
[1] "NOMOR_KONTRAK"      "DOMISILI"  
[3] "KARAKTER"           "PROFESI"  
[5] "KONDISI_USAHA"      "KONDISI_JAMINAN"  
[7] "STATUS"             "PRODUK"  
[9] "PYD"                "TENOR"  
[11] "OSL"                "KEWAJIBAN"  
[13] "KOLEKTIBILITAS"     "COUNT_SURVEY"  
[15] "REKOMENDASI_TINDAK_LANJUT"
```

DATA CONVERSION

Seperti yang diketahui ketika data ditarik dari suatu sumber terkadang ada kondisi tipe data tidak dengan tepat direpresentasikan. Misalkan semua record/baris pada suatu kolom berisi seharusnya data numerik akan tetapi disajikan didalam suatu karakter angka.

R sendiri memiliki fungsi `sapply` yang dapat digunakan untuk mengkonversi tipe data. Dalam hal ini fungsi `sapply` menerima input/argumen fungsi berupa list, vector, atau data frame dan mengembalikan/menghasilkan output berupa vector atau matrix.

Jika tidak perlu di konversi, tidak perlu diubah

Tetapi, jika perlu dirubah kolom "PRODUK", "PYD", "TENOR", dan "OSL" maka perintahnya berikut:

```
data_reduce[, 8:11] = sapply(data_reduce[, 8:11], as.numeric)
```

CATEGORY SELECTION OF DATA

Data kategori dapat dipilih melalui kolom-kolom "KONDISI_USAHA", "KONDISI_JAMINAN", "REKOMENDASI_TINDAK_LANJUT".

Ubah kolom "REKOMENDASI_TINDAK_LANJUT" sebagai faktor (gunakan `as.factor`).

Gunakan uji chi-square dapat digunakan untuk melihat hubungan antar variabel kategorik berikut:

```
1 data_kategorik =  
  data_reduce[,c("KONDISI_USAHA", "KONDISI_JAMINAN", "REKOMENDASI_TINDAK_LANJUT")]  
2 data_reduce$REKOMENDASI_TINDAK_LANJUT =  
  as.factor(data_reduce$REKOMENDASI_TINDAK_LANJUT)  
3 chisq.test(data_kategorik$KONDISI_USAHA, data_kategorik$REKOMENDASI_TINDAK_LANJUT)  
4 chisq.test(data_kategorik$KONDISI_JAMINAN,  
  data_kategorik$REKOMENDASI_TINDAK_LANJUT)
```

Output after input syntax above:

```
> data_kategorik = data_reduce[,c("KONDISI_USAHA", "KONDISI_JAMINAN", "REKOMENDASI_TINDAK_LANJUT")]

> data_reduce$REKOMENDASI_TINDAK_LANJUT = as.factor(data_reduce$REKOMENDASI_TINDAK_LANJUT)

> chisq.test(data_kategorik$KONDISI_USAHA, data_kategorik$REKOMENDASI_TINDAK_LANJUT)

        Pearson's Chi-squared test

data:  data_kategorik$KONDISI_USAHA and data_kategorik$REKOMENDASI_TINDAK_LANJUT
X-squared = 129.82, df = 6, p-value < 2.2e-16

> chisq.test(data_kategorik$KONDISI_JAMINAN, data_kategorik$REKOMENDASI_TINDAK_LANJUT)

        Pearson's Chi-squared test

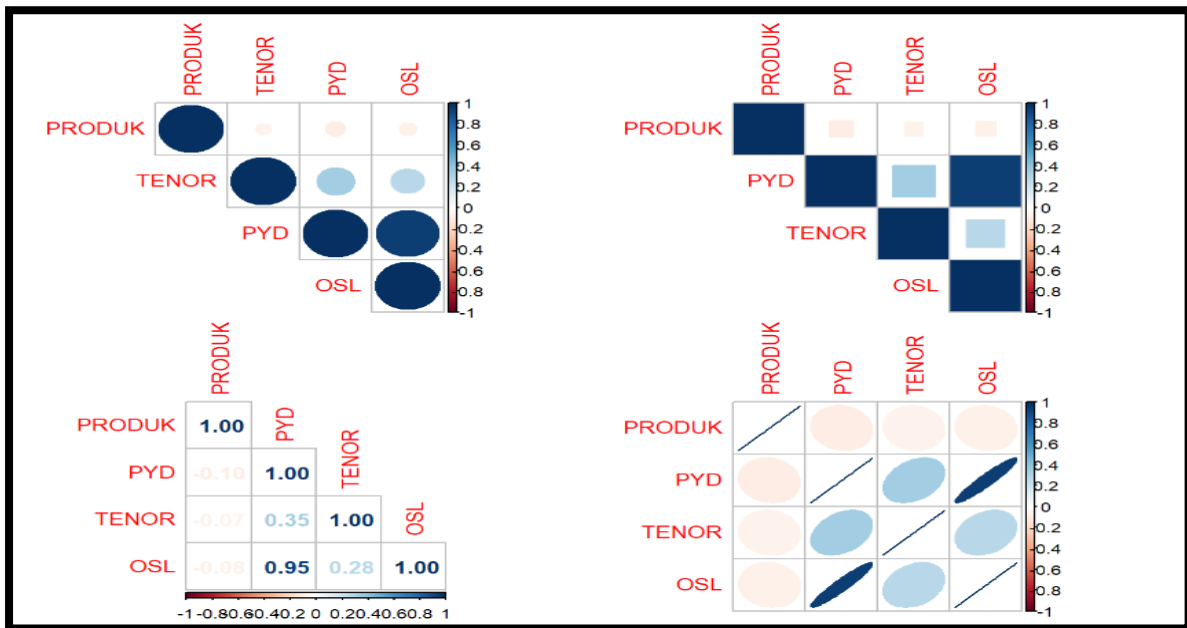
data:  data_kategorik$KONDISI_JAMINAN and data_kategorik$REKOMENDASI_TINDAK_LANJUT
X-squared = 162.87, df = 9, p-value < 2.2e-16
```

CORRELATION BETWEEN DATA VARIABLES

Selain melihat hubungan pada data yang bersifat kategorikal, kita juga bisa melihat hubungan antar variabel numerikal. Ya. Kita akan menggunakan korelasi.

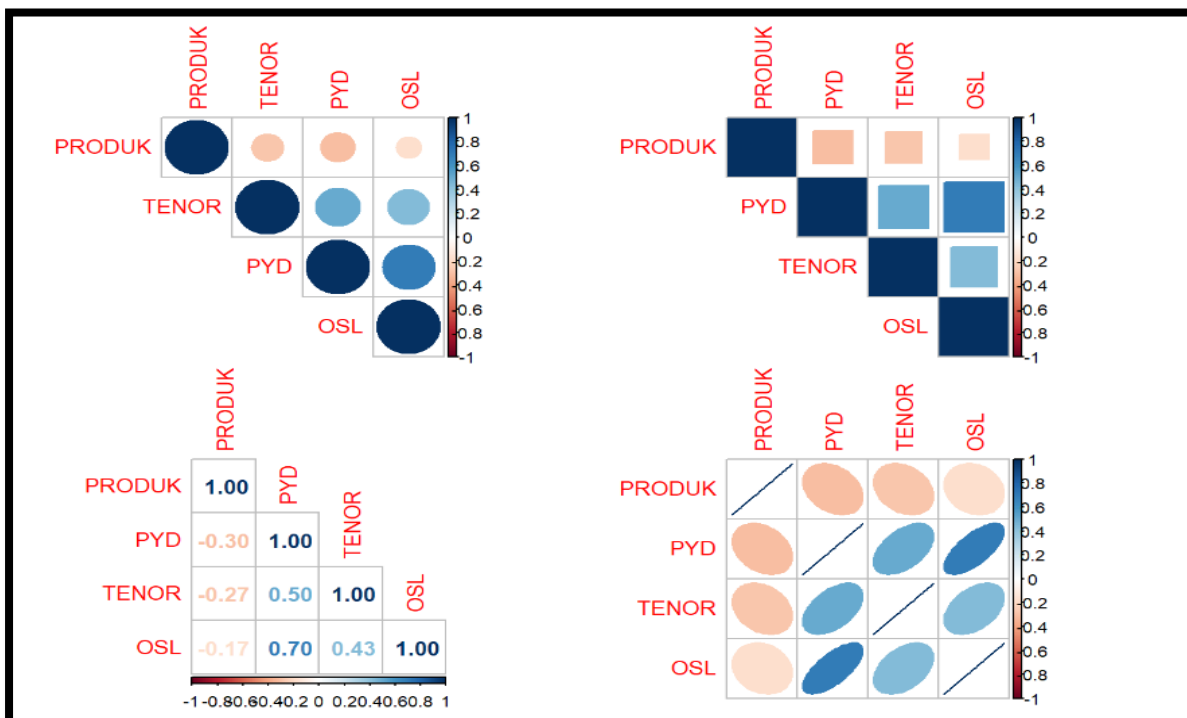
```
1 library(corrplot)
2 library(ggcorrplot)
3
4 M = data_reduce[,8:11]
5
6 # Library corrplot
7 # -- Pearson correlation
8 par(mfrow=c(2,2))
9 corrplot(cor(M), type="upper", order="hclust")
10 corrplot(cor(M), method="square", type="upper")
11 corrplot(cor(M), method="number", type="lower")
12 corrplot(cor(M), method="ellipse")
13
```

Output after input syntax above (Pearson Correlation):



```
14 # -- Kendall correlation
15 par(mfrow=c(2,2))
16 corrplot(cor(M, method="kendall"), type="upper", order="hclust")
17 corrplot(cor(M, method="kendall"), method="square", type="upper")
18 corrplot(cor(M, method="kendall"), method="number", type="lower")
19 corrplot(cor(M, method="kendall"), method="ellipse")
20
```

Output after input syntax above (Kendall Correlation):

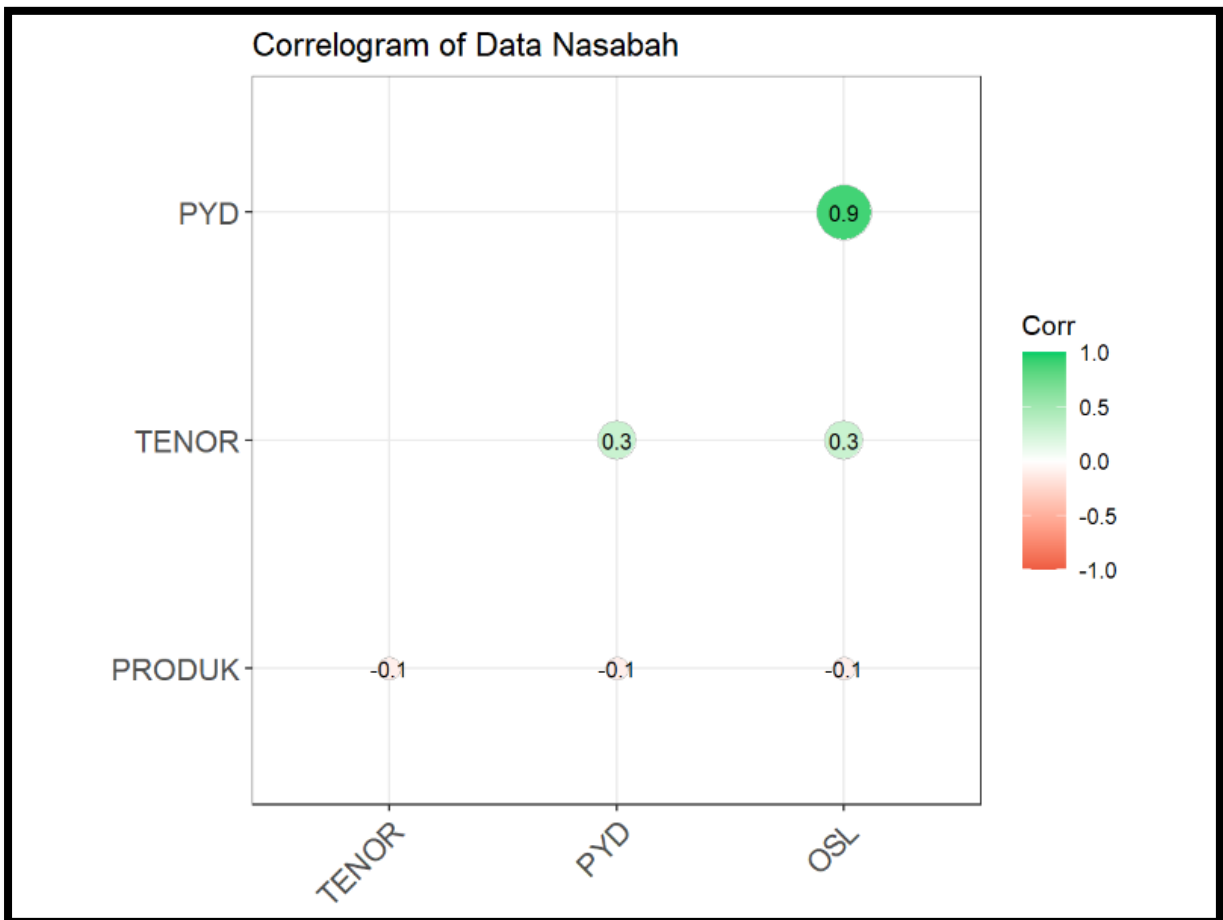


```

21 # Library ggcorrplot
22 corr = round(cor(M), 1) # Pearson correlation
23 ggcorrplot(round(cor(M), 1),
24             hc.order = TRUE,
25             type = "lower",
26             lab = TRUE,
27             lab_size = 3,
28             method="circle",
29             colors = c("tomato2", "white", "springgreen3"),
30             title="Correlogram of Data Nasabah",
31             ggtheme=theme_bw)

```

Output after input syntax above (Correlogram):



Metode yang digunakan untuk melihat korelasi adalah dengan metode korelasi pearson dan korelasi kendall. Jika nilai korelasinya bernilai positif maka korelasi antar variabel tersebut berkorelasi kuat.

FEATURE SELECTION / INDEPENDENT VARIABLE / INPUT

Dalam melakukan pemodelan tentu kita perlu meninjau variabel-variabel apa saja yang berpengaruh pada model kita, khususnya pada klasifikasi. Pada kesempatan ini kita menggunakan model Regresi Multinomial.

Lalu bagaimana menentukan variabel apa saja yang berpengaruh tersebut?

Ada banyak alternatif, salah satunya ialah Information Gain. Melalui information gain diambil nilai importance variabel yang lebih dari 0.02 (kamu dapat eksplorasi apa yang terjadi apabila kita mengambil nilai yang kurang dari 0.02).

```
1 colnames(data_reduce)
2
3 data_select = data_reduce[,
  c("KARAKTER", "KONDISI_USAHA", "KONDISI_JAMINAN", "STATUS", "KEWAJIBAN", "
  OSL", "KOLEKTIBILITAS", "REKOMENDASI_TINDAK_LANJUT")]
4
5 data_non_na = na.omit(data_select)
```

Output after input syntax above:

```
> colnames(data_reduce)
[1] "NOMOR_KONTRAK"      "DOMISILI"
[3] "KARAKTER"           "PROFESI"
[5] "KONDISI_USAHA"      "KONDISI_JAMINAN"
[7] "STATUS"             "PRODUK"
[9] "PYD"               "TENOR"
[11] "OSL"                "KEWAJIBAN"
[13] "KOLEKTIBILITAS"     "COUNT_SURVEY"
[15] "REKOMENDASI_TINDAK_LANJUT"

> data_select = data_reduce[, c("KARAKTER", "KONDISI_USAHA", "KONDISI_JAMINAN", "STATUS", "KEWAJIBAN", "OSL", "KOLEKTIBILITAS", "REKOMENDASI_TINDAK_LANJUT")]

> data_non_na = na.omit(data_select)
```

DATA TRANSFORMATION

Untuk memberikan performa model yang baik, maka pada data kita perlu dilakukan treatment tertentu, misalnya dilakukan scaling atau dilakukan pengelompokan data atau disebut juga *bucketing*.

```
1 data_select_new = data_select
2 data_select_new$KEWAJIBAN = scale(data_select_new$KEWAJIBAN)[, 1]
3 data_select_new$OSL = scale(data_select_new$OSL)[, 1]
4 data_select_new$KEWAJIBAN = cut(data_select_new$KEWAJIBAN, breaks =
  c(-0.354107, 5, 15, 30))
5 data_select_new$KEWAJIBAN = as.factor(data_select_new$KEWAJIBAN)
6 data_select_new$OSL = cut(data_select_new$OSL, breaks = c(-0.60383, 3, 10, 15))
7 data_select_new$OSL = as.factor(data_select_new$OSL)
8 data_select_new = na.omit(data_select_new)
```


Output after input syntax above:

```
> data_select_new = data_select
> data_select_new$KEWAJIBAN = scale(data_select_new$KEWAJIBAN)[, 1]
> data_select_new$OSL = scale(data_select_new$OSL)[, 1]
> data_select_new$KEWAJIBAN = cut(data_select_new$KEWAJIBAN, breaks = c(-0.354107,5,15,30))
> data_select_new$KEWAJIBAN = as.factor(data_select_new$KEWAJIBAN)
> data_select_new$OSL = cut(data_select_new$OSL, breaks = c(-0.60383,3,10,15))
> data_select_new$OSL = as.factor(data_select_new$OSL)
> data_select_new = na.omit(data_select_new)
```

TRAINING DATA

```
1 library(caret)
2 index = createDataPartition(data_select_new$REKOMENDASI_TINDAK_LANJUT, p = .95,
  list = FALSE)
3 train = data_select_new[index,]
4 test = data_select_new[-index,]
```

Output after input syntax above:

```
> library(caret)
> index = createDataPartition(data_select_new$REKOMENDASI_TINDAK_LANJUT, p = .95, list = FALSE)
> train = data_select_new[index,]
> test = data_select_new[-index,]
```

MODELLING

Let's modelling the data

```
1 train2 = train
2 # Setting the reference
3 train2$REKOMENDASI_TINDAK_LANJUT = relevel(train2$REKOMENDASI_TINDAK_LANJUT, ref = "Angsuran Biasa")
4 #training the model
5 require(nnet)
6 # Training the multinomial model
7 multinom_model = multinom(REKOMENDASI_TINDAK_LANJUT ~ ., data = train2)
8
```

```

9 # Checking the model
10 summary(multinom_model)
11 #converting the coefficients to odds by taking the exponential of the
    coefficients.
12 exp(coef(multinom_model))
13 head(round(fitted(multinom_model), 2))
14 # Predicting the values for train dataset
15 train2$ClassPredicted = predict(multinom_model, newdata = train2, "class")
16 train_prob = predict(multinom_model, newdata = train2, "probs")
17 df = train_prob
18 df$max = apply(df,1, max)
19 train2$score = df$max
20 test_prob = predict(multinom_model, newdata = test, "probs")
21 df2 = test_prob
22 df2$max = apply(df2,1, max)
23
24 # Building classification table
25 tab_train = table(train2$REKOMENDASI_TINDAK_LANJUT, train2$ClassPredicted)
26 round((sum(diag(tab_train))/sum(tab_train))*100,4)
27 test$ClassPredicted = predict(multinom_model, newdata = test, "class")
28 test$score = df2$max
29 tab_test = table(test$REKOMENDASI_TINDAK_LANJUT, test$ClassPredicted)
30 round((sum(diag(tab_test))/sum(tab_test))*100,4)
31

```

Output after input syntax above:

```

> train2 = train

> # Setting the reference
> train2$REKOMENDASI_TINDAK_LANJUT = relevel(train2$REKOMENDASI_TINDAK_LANJUT, ref = "Angs
uran Biasa")

> #training the model
> require(nnet)

> # Training the multinomial model
> multinom_model = multinom(REKOMENDASI_TINDAK_LANJUT ~ ., data = train2)
# weights: 64 (45 variable)
initial value 1319.752232
iter 10 value 746.560598
iter 20 value 617.979680
iter 30 value 613.601197
iter 40 value 613.309770
iter 50 value 613.301463
iter 60 value 613.300961
final value 613.300951
converged

> # Checking the model
> summary(multinom_model)
Call:
multinom(formula = REKOMENDASI_TINDAK_LANJUT ~ ., data = train2)

Coefficients:
              (Intercept) KARAKTERTIDAK KOOPERATIF KONDISI_USAHA
Diskon Pelunasan    -3.289036              1.782955             -0.7320989

```

```

Penarikan -7.059111 3.775381 0.6189923
Restrukturisasi -4.169442 -2.260618 1.0165643
KONDISI_JAMINANHilang KONDISI_JAMINANPindah Tangan
Diskon Pelunasan -25.3254192 -29.1228976
Penarikan 0.8744536 0.5602014
Restrukturisasi 0.4645554 -13.6439503
KONDISI_JAMINANRusak STATUS KEWAJIBAN(5,15]
Diskon Pelunasan -26.9472468 0.09492752 0
Penarikan -0.9168574 0.11590743 0
Restrukturisasi 0.1132984 0.31876307 0
KEWAJIBAN(15,30] OSL(3,10] OSL(10,15]
Diskon Pelunasan -4.079239 2.9877112 0
Penarikan 1.110055 -26.4876072 0
Restrukturisasi 28.435937 0.4292003 0
KOLEKTIBILITASDIRAGUKAN KOLEKTIBILITASKURANG LANCAR
Diskon Pelunasan 2.7351334 0.2863512
Penarikan 3.9267484 0.9129418
Restrukturisasi -0.3026627 -0.1927960
KOLEKTIBILITASLANCAR KOLEKTIBILITASMACET
Diskon Pelunasan 0.1771182 0.653555
Penarikan -33.0132242 1.551052
Restrukturisasi 0.8696888 -3.172181

Std. Errors:
(Intercept) KARAKTERTIDAK KOOPERATIF KONDISI_USAHA
Diskon Pelunasan 0.4029650 0.05385465 NaN
Penarikan 0.9473535 0.75505818 NaN
Restrukturisasi NaN NaN NaN
KONDISI_JAMINANHilang KONDISI_JAMINANPindah Tangan
Diskon Pelunasan 0.08568768 NaN
Penarikan 0.73958528 0.9155759
Restrukturisasi NaN NaN
KONDISI_JAMINANRusak STATUS KEWAJIBAN(5,15] KEWAJIBAN(15,30]
Diskon Pelunasan NaN NaN NaN NaN
Penarikan 0.8782296 NaN 0.8805277 NaN
Restrukturisasi NaN 0.689277 0.0000000 NaN
OSL(3,10] OSL(10,15] KOLEKTIBILITASDIRAGUKAN
Diskon Pelunasan 0.3246491 NaN 0.9961205
Penarikan 1.2791632 NaN 1.0858551
Restrukturisasi NaN 0 0.1275660
KOLEKTIBILITASKURANG LANCAR KOLEKTIBILITASLANCAR
Diskon Pelunasan 0.5803431 0.62590746
Penarikan 1.0811309 NaN
Restrukturisasi NaN 0.07679056
KOLEKTIBILITASMACET
Diskon Pelunasan NaN
Penarikan 1.010654
Restrukturisasi NaN

```

Residual Deviance: 1226.602
AIC: 1304.602

```

> #converting the coefficients to odds by taking the exponential of the coefficients.
> exp(coef(multinom_model))
(Intercept) KARAKTERTIDAK KOOPERATIF KONDISI_USAHA
Diskon Pelunasan 0.0372897767 5.947402 0.4808986
Penarikan 0.0008595417 43.614108 1.8570557
Restrukturisasi 0.0154608884 0.104286 2.7636833
KONDISI_JAMINANHilang KONDISI_JAMINANPindah Tangan

```

```

Diskon Pelunasan      1.003021e-11      2.249501e-13
Penarikan             2.397565e+00      1.751025e+00
Restrukturisasi      1.591307e+00      1.187156e-06
      KONDISI_JAMINANRusak      STATUS KEWAJIBAN(5,15] KEWAJIBAN(15,30]
Diskon Pelunasan      1.981342e-12  1.099579      1      1.692034e-02
Penarikan             3.997734e-01  1.122892      1      3.034526e+00
Restrukturisasi      1.119966e+00  1.375425      1      2.236507e+12
      OSL(3,10] OSL(10,15] KOLEKTIBILITASDIRAGUKAN
Diskon Pelunasan      1.984022e+01      1      15.4117996
Penarikan             3.137461e-12      1      50.7417180
Restrukturisasi      1.536029e+00      1      0.7388482
      KOLEKTIBILITASKURANG LANCAR KOLEKTIBILITASLANCAR
Diskon Pelunasan      1.3315600      1.193772e+00
Penarikan             2.4916418      4.597682e-15
Restrukturisasi      0.8246502      2.386168e+00
      KOLEKTIBILITASMACET
Diskon Pelunasan      1.92236263
Penarikan             4.71642843
Restrukturisasi      0.04191209

> head(round(fitted(multinom_model), 2))
  Angsuran Biasa Diskon Pelunasan Penarikan Restrukturisasi
1          0.81          0.01          0.00          0.18
2          0.46          0.00          0.53          0.01
3          0.25          0.03          0.72          0.00
4          0.98          0.00          0.01          0.02
5          0.44          0.19          0.37          0.00
6          0.70          0.02          0.00          0.28

> # Predicting the values for train dataset
> train2$ClassPredicted = predict(multinom_model, newdata = train2, "class")

> train_prob = predict(multinom_model, newdata = train2, "probs")

> df = train_prob

> df$max = apply(df,1, max)

> train2$score = df$max

> test_prob = predict(multinom_model, newdata = test, "probs")

> df2 = test_prob

> df2$max = apply(df2,1, max)

> # Building classification table
> tab_train = table(train2$REKOMENDASI_TINDAK_LANJUT, train2$ClassPredicted)

> round((sum(diag(tab_train))/sum(tab_train))*100,4)
[1] 68.9076

> test$ClassPredicted = predict(multinom_model, newdata = test, "class")
> test$score = df2$max
> tab_test = table(test$REKOMENDASI_TINDAK_LANJUT, test$ClassPredicted)
> round((sum(diag(tab_test))/sum(tab_test))*100,4)
[1] 64.5833

```

Berdasarkan hasil di atas, maka didapatkan:

- Akurasi model untuk training sebesar : 68,91%
- Akurasi model untuk testing sebesar : 64,58%

CLOSING

That's all for this project and thanks for watching.

GET IN TOUCH :

LinkedIn [linkedin.com/in/vicky-jodie](https://www.linkedin.com/in/vicky-jodie)

GitHub github.com/vickyjodie