Laporan Tugas 5

Analisis Klasifikasi dan Korelasi Kanonik

Analisis Multivariat Kelas B

Kelompok L

Evan Haryowidyatna	2006485011
Muhammad Jauhar Hakim	2006463982
Siskawati Simandalahi	2006572970

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Indonesia

Depok

April 2022

I. Penjelasan Data

1.1 Analisis Klasifikasi

Data yang kami peroleh merupakan data kualitas wine. Kami mendapatkan data dari situs https://archive.ics.uci.edu/ml/datasets/Wine+Quality Data kualitas wine terdiri dari 12 atribut yaitu:

1. Fixed acidity (keasaman tetap)
2. Volatile acidity (keasaman yang menguap)
3. Citric acid (asam sitrat)
4. Residual sugar (sisa gula)
5. Chlorides (klorida)
6. Free sulfur dioxide (sulfur dioksida bebas)
7. Total sulfur dioxide (sulfur dioksida total)
8. Density (kepadatan)
9. pH
10. Sulphates (sulfat)
11. Alcohol (alkohol)
12. Quality (score between 0 and 10)

1.2 Korelasi Kanonik

Data yang kami peroleh merupakan data kualitas udara. Kami mendapatkan data dari situs https://archive.ics.uci.edu/ml/datasets/air+quality Data kualitas udara terdiri dari 15 atribut yaitu:

1. Date (DD/MM/YYYY)
2. Time (HH.MM.SS)
3. True hourly averaged concentration CO in mg/m^3 (reference analyzer)
4. PT08.S1 (tin oxide) hourly averaged sensor response (nominally CO targeted)
5. True hourly averaged overall Non Metanic HydroCarbons concentration in microg/m^3 (reference analyzer)
6. True hourly averaged Benzene concentration in microg/m^3 (reference analyzer)
7. PT08.S2 (titania) hourly averaged sensor response (nominally NMHC targeted)
8. True hourly averaged NOx concentration in ppb (reference analyzer)
9. PT08.S3 (tungsten oxide) hourly averaged sensor response (nominally NOx targeted)

10. True hourly averaged NO2 concentration in microg/m^3 (reference analyzer)
11. PT08.S4 (tungsten oxide) hourly averaged sensor response (nominally NO2 targeted)
12. PT08.S5 (indium oxide) hourly averaged sensor response (nominally O3 targeted)
13. Temperature in Â°C
14. Relative Humidity (%)
15. AH Absolute Humidity

II. Tujuan

2.1 Analisis Klasifikasi

Analisis klasifikasi adalah metode untuk menganalisis keterkaitan antara beberapa variabel independen dan satu variabel dependen yang merupakan variabel kualitatif. Beberapa variabel prediktor ini akan digunakan untuk memprediksi kategori atau kelas suatu variabel dependen. Metode yang digunakan untuk pengklasifikasian pertama adalah memprediksi peluang dari setiap kategori dari variabel kualitatif sebagai dasar untuk membuat klasifikasi (James et al, 2013).

2.2 Korelasi Kanonik

Analisis korelasi kanonik merupakan suatu teknik yang berguna untuk mengidentifikasi dan mengukur hubungan linier, yang melibatkan beberapa variabel dependen dan independen. Korelasi kanonik fokus pada korelasi antara kombinasi linier dari suatu himpunan variabel independen dengan kombinasi linier dari himpunan variabel dependen. Pasangan kombinasi liniernya disebut fungsi kanonik, dan korelasinya disebut koefisien korelasi kanonik.

III. Dasar Teori

3.1 Analisis Klasifikasi

Klasifikasi merupakan proses untuk menemukan sekumpulan model yang menjelaskan dan membedakan kelas-kelas data, sehingga model tersebut dapat digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek. Proses klasifikasi data didasarkan oleh 4 komponen mendasar antara lain yaitu kelas, prediktor, training set, serta pengujian dataset.

Sejumlah teknik klasifikasi telah banyak diusulkan dan dikembangkan dalam literatur. Terutama proses klasifikasi dibagi menjadi beberapa kategori yang berbeda, yang dinamakan sebagai keputusan berbasis pengklasifikasi. Beberapa metode klasifikasi tersebut diantaranya:

1. Decision Tree
   Decision Tree mewakili serangkaian keputusan dan pilihan dalam bentuk pohon. Decision Tree menggunakan fitur-fitur dari suatu objek untuk memutuskan kelas di mana objek itu berada. Kelas-kelas ini biasanya terletak di terminal leaver dari decision tree dan dapat berupa pengklasifikasi biner atau multi-kelas.

2. Naïve Bayes
   Naïve Bayes merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes . Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes , yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes . Ciri utama dr Naïve Bayes Classifier ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian.

3. Logistic Regression
   Logistic Regression adalah sebuah algoritma klasifikasi untuk mencari hubungan antara fitur (input) diskrit/kontinu dengan probabilitas hasil output diskrit tertentu. Terdapat tiga tipe-tipe logistic regression yaitu Binary Logistic Regression, Multinomial Logistic Regression, dan Ordinal Logistic Regression.

4. K-Nearest Neighbour
   K-Nearest Neighbour adalah algoritma klasifikasi sederhana yang menyimpan semua kasus yang tersedia dan mengklasifikasikan kasus baru berdasarkan ukuran kesamaan (misalnya, fungsi jarak). KNN telah digunakan dalam estimasi statistik dan pengenalan pola pada awal tahun 1970-an sebagai teknik non-parametrik. Jarak yang digunakan bisa berupa jarak euclidean, jarak manhattan, ataupun jarak minkowski.

5. Support Vector Machine (SVM)
   SVM digunakan untuk mencari hyperplane terbaik dengan memaksimalkan jarak antar kelas. Hyperplane adalah sebuah fungsi yang dapat digunakan untuk pemisah antar kelas. Dalam 2-D fungsi yang digunakan untuk klasifikasi antar kelas disebut sebagai line whereas, fungsi yang digunakan untuk klasifikasi antas kelas dalam 3-D disebut plane similarly, sedangan fungsi yang digunakan untuk klasifikasi di dalam ruang kelas dimensi yang lebih tinggi di sebut hyperplane.

6. Artificial Neural Network
   Artificial Neural Network (ANN) atau jaringan syaraf tiruan adalah jaringan dari sekelompok unit pemroses kecil yang dimodelkan berdasarkan perilaku jaringan syaraf manusia.

3.2 Korelasi kanonik

Analisis korelasi kanonik pertama kali diperkenalkan oleh Hotelling pada tahun 1936, sebagai suatu teknik statistika peubah ganda yang menyelediki keeratan hubungan antara dua gugus peubah. Analisis korelasi kanonikal adalah model statistika multivariat yang memungkinkan identifikasi dan kuantifikasi hubungan antara dua himpunan variabel (Hair, Anderson, Tatham, & Black, 2010).

Menurut (Irianingsih, Gusriani, Kulsum, & Parmikanti, n.d.) Analisis korelasi kanonik merupakan teknik multivariat yang digunakan untuk mengestimasi hubungan antara dua atau lebih variabel dependen dengan dua atau lebih variabel independen secara bersama-sama. Dengan korelasi kanonikal (canonical correlation), dapat menghubungkan beberapa variabel dependen dengan beberapa variabel independen sekaligus. Langkah pertama adalah mencari kombinasi linier yang memiliki korelasi terbesar. Selanjutnya, akan dicari pasangan kombinasi linier dengan nilai korelasi terbesar di antara semua pasangan lain yang tidak berkorelasi. Proses terjadi secara berulang, hingga korelasi maksimum teridentifikasi. Pasangan kombinasi linier disebut sebagai variat kanonikal sedangkan hubungan di antara pasangan tersebut disebut korelasi kanonikal.

Cara menghitung r adalah sebagai berikut :

$$Koefisien\ korelasi\ (r) = \frac{n(\sum X_i Y_i) - (\sum X_i)(\sum Y_i)}{\sqrt{n\ ((\sum X_i^2)\ (\sum X_i)^2\ nn(\sum Y_i^2)\ \ (\sum Y_i)^2}}$$

Menurut Young, dalam (Wahid, 2004), ukuran korelasi dinyatakan sebagai berikut:

a) Nilai korelasi 0,7 sampai 1,0 (baik positif maupun negatif) menunjukkan adanya tingkat hubungan yang tinggi.

b) Nilai korelasi 0,4 sampai < 0,7 (baik positif maupun negatif) menunjukkan adanya tingkat hubungan yang substansial.

c) Nilai korelasi 0,2 sampai < 0,4 (baik positif maupun negatif) menunjukkan adanya tingkat hubungan yang rendah.

d) Nilai korelasi < 0,2 (baik positif maupun negatif) menunjukkan tidak adanya tingkat hubungan

IV. Pemodelan

4.1 Analisis Klasifikasi

Berikut ini adalah langkah kerja untuk melakukan analisis klasifikasi:

1. Mengimpor data kualitas wine
2. Melakukan uji asumsi multivariat normal
3. Mendefiniskan kategori untuk kualitas wine
4. Klasifikasi menggunakan *decision tree*
5. Klasifikasi menggunakan kNN *(k-Nearest Neighbors)*

4.2 Korelasi Kanonik

Berikut ini adalah langkah kerja untuk analisis korelasi kanonik:

1. Mengimpor data kualitas udara
2. Mendefinisikan variabel dependen dan variabel independen
3. Mencari korelasi kanonik

IV. Proses Komputasi

5.1 Analisis Klasifikasi

Install package dan *apply library*

```
packages <- c("Hmisc", "matlib",
"Matrix","expm","matrixcalc","ellipsis","Hotelling","dplyr","psych","Rc
mdrMisc","Rcsdp","mvnormtest","factoextra","cluster","ggplot2","tree","
class")

if ( length(missing_pkgs <- setdiff(packages,
rownames(installed.packages()))) > 0) {

  message("Installing missing package(s): ", paste(missing_pkgs,
collapse = ", "))

  install.packages(missing_pkgs)

}

lapply(packages, library, character.only = TRUE)
```

1. Mengimpor data kualitas wine

```
data <- read.table("winequality-red.csv", header=TRUE, sep=";")

data <- as.data.frame(data)

data
```

2. Melakukan uji asumsi multivariat normal

```
mshapiro.test(t(data))
```

3. Mendefinisikan kategori untuk kualitas wine

```
data$type <- as.factor(ifelse(data$quality <= 5, 'Kurang Baik', 'Cukup
Baik'))

data

str(data)
```

4. Klasifikasi menggunakan *decision tree*

```
quality.tree <- tree(type~.-quality,data = data)
summary(quality.tree)
options(repr.plot.width = 13, repr.plot.height = 7, repr.plot.res = 100)
plot(quality.tree)
text(quality.tree, pretty=0)

set.seed(40)
tree.train <- sample(1:nrow(data),250)
quality.tree <- tree(type~.-quality,data,subset=tree.train)
options(repr.plot.width = 13, repr.plot.height = 7, repr.plot.res = 100)
plot(quality.tree)
text(quality.tree, pretty=0)

quality.pred = predict(quality.tree, data[-tree.train,], type="class")
with(data[-tree.train,], table(quality.pred, type))

quality.cv = cv.tree(quality.tree, FUN = prune.misclass)
quality.cv

plot(quality.cv)
quality.prune = prune.misclass(quality.tree, best = 12)
plot(quality.prune)
text(quality.prune, pretty=0)

quality.pred = predict(quality.prune, data[-tree.train,], type="class")
with(data[-tree.train,], table(quality.pred, type))
```

5. Klasifikasi menggunakan kNN *(k-Nearest Neighbors)*

```
# Count the number of signs of each type
table(data$type)
# Use kNN to identify the test road signs
data_types <- data$type
data_pred <- knn(train = data[-13], test = data[-13], cl = data_types)
# Create a confusion matrix of the predicted versus actual values
data_actual <- data$type
table(data_pred, data_actual)


# Compute the accuracy
mean(data_pred == data_actual)
# Compute the accuracy of the baseline model (default k = 1)
k_1 <- knn(train = data[-13], test = data[-13], cl = data_types)
mean(k_1 == data_actual)
# Modify the above to set k = 7
k_7 <- knn(train = data[-13], test = data[-13], cl = data_types, k = 7)
mean(k_7 == data_actual)
# Use the prob parameter to get the proportion of votes for the winning class
data_pred <- knn(train = data[-13], test = data[-13], cl = data_types,
k = 7, prob = TRUE)

# Get the "prob" attribute from the predicted classes
data_prob <- attr(data_pred, "prob")
# Examine the first several predictions
head(data_pred)
# Examine the proportion of votes for the winning class
head(data_prob)
```

5.2 Korelasi Kanonik

Install package dan *apply library*

```
packages <- c("Hmisc", "matlib",
"Matrix","expm","matrixcalc","ellipsis","Hotelling","dplyr","psych","Rc
mdrMisc","Rcsdp","mvnormtest","factoextra","cluster","ggplot2","tree","
class","CCA","vegan","candisc")

if ( length(missing_pkgs <- setdiff(packages,
rownames(installed.packages()))) > 0) {

  message("Installing missing package(s): ", paste(missing_pkgs,
collapse = ", "))

  install.packages(missing_pkgs)

}

lapply(packages, library, character.only = TRUE)
```

1. Mengimpor data kualitas udara

```
data2 <- read.csv("AirQualityUCI.csv", header=TRUE, sep=";")

head(data2)

databaru=data2[-1:-2]

databaru2=databaru[-14:-15]

datafinal = na.omit(databaru2)

head(datafinal)
```

2. Mendefinisikan variabel dependen dan variabel independen

```
X <- datafinal[1:10]

Y <- datafinal[11:13]

head(X)

head(Y)
```

3. Mencari korelasi kanonik

```
library("CCA")

correl <- matcor(X, Y )

correl

img.matcor(correl, type = 2)


cc1 <- cancor(X, Y)  ### function from standard R instalation

cc2 <- cc(X, Y)      ### function for the R package 'CCA'

cc1

cc2


par(mfrow = c(1,2))

barplot(cc1$cor, main = "Canonical correlations for 'cancor()'", col =
"gray")

barplot(cc1$cor, main = "Canonical correlations for 'cancor()'", col =
"gray")

cc1$xcoef  ### function from standard R instalation

plt.cc(cc2, var.label = TRUE)


# ANALISIS KORELASI KANONIK

ccan <- candisc::cancor(X,Y)

summary(ccan)

# Korelasi kanonik

res.cc <- cc(X,Y)

res.cc

# Plot korelasi

plot(res.cc$cor,type="b")
```

V. Hasil Komputasi

6.1 Analisis Klasifikasi

1. Mengimpor data kualitas wine

fixed.acidity	volatile.acidity	citric.acid	residual.sugar	chlorides	free.sulfur.dioxide	total.s
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
7.4	0.700	0.00	1.9	0.076	11	34
7.8	0.880	0.00	2.6	0.098	25	67
7.8	0.760	0.04	2.3	0.092	15	54
11.2	0.280	0.56	1.9	0.075	17	60
7.4	0.700	0.00	1.9	0.076	11	34
7.4	0.660	0.00	1.8	0.075	13	40
7.9	0.600	0.06	1.6	0.069	15	59
7.3	0.650	0.00	1.2	0.065	15	21
7.8	0.580	0.02	2.0	0.073	9	18
7.5	0.500	0.36	6.1	0.071	17	102
6.7	0.580	0.08	1.8	0.097	15	65
7.5	0.500	0.36	6.1	0.071	17	102
5.6	0.615	0.00	1.6	0.089	16	59
7.8	0.610	0.29	1.6	0.114	9	29
8.9	0.620	0.18	3.8	0.176	52	145
8.9	0.620	0.19	3.9	0.170	51	148
8.5	0.280	0.56	1.8	0.092	35	103
8.1	0.560	0.28	1.7	0.368	16	56
7.4	0.590	0.08	4.4	0.086	6	29

2. Melakukan uji asumsi multivariat normal

```
        Shapiro-Wilk normality test

data:  Z
W = 0.62625, p-value < 2.2e-16
```

3. Mendefinisikan kategori untuk kualitas wine

free.sulfur.dioxide	total.sulfur.dioxide	density	pH	sulphates	alcohol	quality	type
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	<fct>
11	34	0.9978	3.51	0.56	9.4	5	Kurang Baik
25	67	0.9968	3.20	0.68	9.8	5	Kurang Baik
15	54	0.9970	3.26	0.65	9.8	5	Kurang Baik
17	60	0.9980	3.16	0.58	9.8	6	Cukup Baik
11	34	0.9978	3.51	0.56	9.4	5	Kurang Baik
13	40	0.9978	3.51	0.56	9.4	5	Kurang Baik
15	59	0.9964	3.30	0.46	9.4	5	Kurang Baik
15	21	0.9946	3.39	0.47	10.0	7	Cukup Baik
9	18	0.9968	3.36	0.57	9.5	7	Cukup Baik
17	102	0.9978	3.35	0.80	10.5	5	Kurang Baik
15	65	0.9959	3.28	0.54	9.2	5	Kurang Baik
17	102	0.9978	3.35	0.80	10.5	5	Kurang Baik
16	59	0.9943	3.58	0.52	9.9	5	Kurang Baik
9	29	0.9974	3.26	1.56	9.1	5	Kurang Baik
52	145	0.9986	3.16	0.88	9.2	5	Kurang Baik
51	148	0.9986	3.17	0.93	9.2	5	Kurang Baik
35	103	0.9969	3.30	0.75	10.5	7	Cukup Baik
16	56	0.9968	3.11	1.28	9.3	5	Kurang Baik
6	29	0.9974	3.38	0.50	9.0	4	Kurang Baik

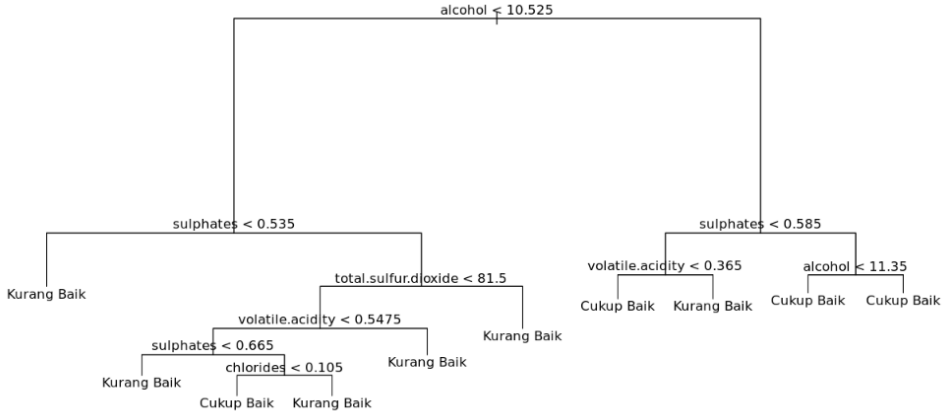
< 5 = Kurang baik

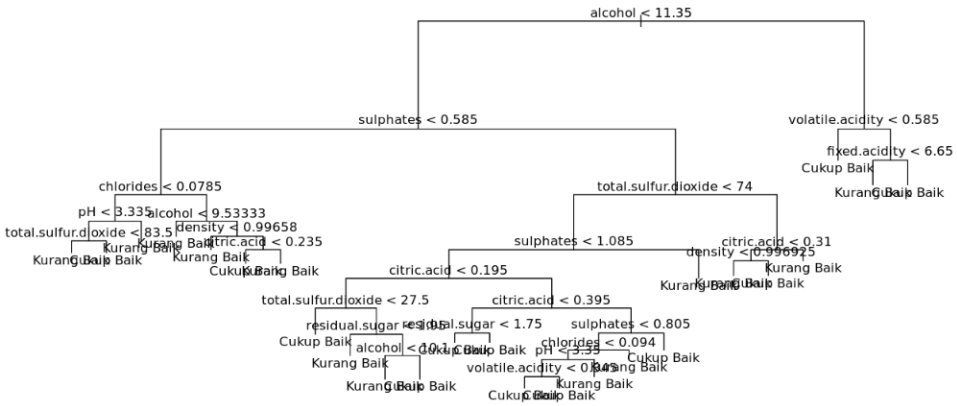
> 5 = Cukup baik

```
'data.frame':   1599 obs. of  13 variables:
 $ fixed.acidity       : num  7.4 7.8 7.8 11.2 7.4 7.4 7.9 7.3 7.8 7.5 ...
 $ volatile.acidity    : num  0.7 0.88 0.76 0.28 0.7 0.66 0.6 0.65 0.58 0.5 ...
 $ citric.acid         : num  0 0 0.04 0.56 0 0 0.06 0 0.02 0.36 ...
 $ residual.sugar      : num  1.9 2.6 2.3 1.9 1.9 1.8 1.6 1.2 2 6.1 ...
 $ chlorides           : num  0.076 0.098 0.092 0.075 0.076 0.075 0.069 0.065 0.073 0.071 ...
 $ free.sulfur.dioxide : num  11 25 15 17 11 13 15 15 9 17 ...
 $ total.sulfur.dioxide: num  34 67 54 60 34 40 59 21 18 102 ...
 $ density             : num  0.998 0.997 0.997 0.998 0.998 ...
 $ pH                  : num  3.51 3.2 3.26 3.16 3.51 3.51 3.3 3.39 3.36 3.35 ...
 $ sulphates           : num  0.56 0.68 0.65 0.58 0.56 0.56 0.46 0.47 0.57 0.8 ...
 $ alcohol             : num  9.4 9.8 9.8 9.8 9.4 9.4 9.4 10 9.5 10.5 ...
 $ quality             : int  5 5 5 6 5 5 5 7 7 5 ...
 $ type                : Factor w/ 2 levels "Cukup Baik","Kurang Baik": 2 2 2 1 2 2 2 1 1 2 ...
```

4. Klasifikasi menggunakan decision tree

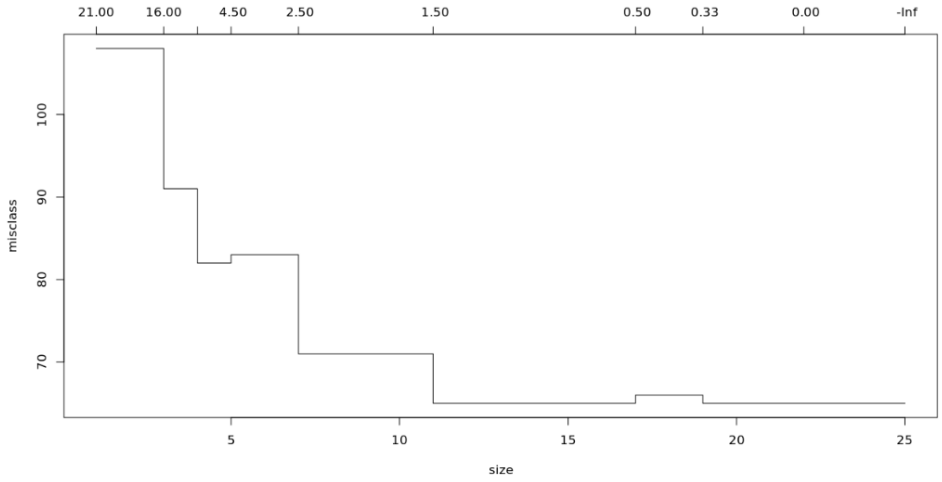
```
Classification tree:
tree(formula = type ~ . - quality, data = data)
Variables actually used in tree construction:
[1] "alcohol"              "sulphates"            "total.sulfur.dioxide"
[4] "volatile.acidity"     "chlorides"
Number of terminal nodes:  10
Residual mean deviance:  1.011 = 1607 / 1589
Misclassification error rate: 0.2539 = 406 / 1599
```

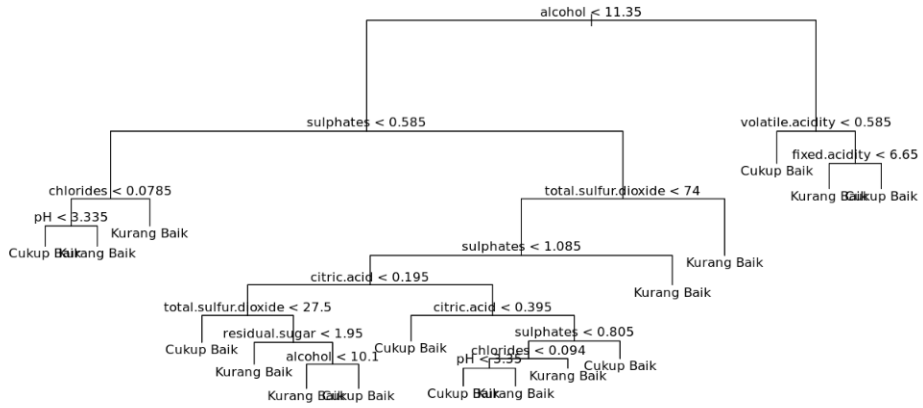




	type	
quality.pred	Cukup Baik	Kurang Baik
Cukup Baik	524	239
Kurang Baik	194	392

```
$size
 [1] 25 22 19 17 11  7  5  4  3  1

$dev
 [1]  65  65  65  66  65  71  83  82  91 108

$k
 [1]       -Inf  0.0000000  0.3333333  0.5000000  1.5000000  2.5000000
 [7]  4.5000000  5.0000000 16.0000000 21.0000000

$method
[1] "misclass"

attr(,"class")
[1] "prune"         "tree.sequence"
```





	type	
quality.pred	Cukup Baik	Kurang Baik
Cukup Baik	520	223
Kurang Baik	198	408

5. Klasifikasi menggunakan kNN *(k-Nearest Neighbors)*

```
# Count the number of signs of each type
table(data$type)
```

```
 Cukup Baik Kurang Baik
        855         744
```

```
# Use kNN to identify the test road signs
data_types <- data$type
data_pred <- knn(train = data[-13], test = data[-13], cl = data_types)

# Create a confusion matrix of the predicted versus actual values
data_actual <- data$type
table(data_pred, data_actual)
```

```
             data_actual
data_pred     Cukup Baik Kurang Baik
  Cukup Baik         855           0
  Kurang Baik          0         744
```

```
# Compute the accuracy
mean(data_pred == data_actual)
```

```
1
```

```
# Compute the accuracy of the baseline model (default k = 1)
k_1 <- knn(train = data[-13], test = data[-13], cl = data_types)
mean(k_1 == data_actual)
```

```
1
```

```
# Modify the above to set k = 7
k_7 <- knn(train = data[-13], test = data[-13], cl = data_types, k = 7)
mean(k_7 == data_actual)
```

```
0.821763602251407
```

```
# Use the prob parameter to get the proportion of votes for the winning class
data_pred <- knn(train = data[-13], test = data[-13], cl = data_types, k = 7, prob = TRUE)


# Get the "prob" attribute from the predicted classes
data_prob <- attr(data_pred, "prob")

# Examine the first several predictions
head(data_pred)
```

```
Kurang Baik · Cukup Baik · Kurang Baik · Cukup Baik · Kurang Baik · Kurang Baik
▼ Levels:
'Cukup Baik' · 'Kurang Baik'
```

```
# Examine the proportion of votes for the winning class
head(data_prob)
```

```
0.714285714285714 · 0.571428571428571 · 1 · 0.571428571428571 · 0.714285714285714 · 0.857142857142857
```

6.2 Analisis Kanonik

1. Mengimpor data kualitas udara

	Date	Time	CO.GT.	PT08.S1.CO.	NMHC.GT.	C6H6.GT.	PT08.S2.NMHC.	NOx.GT.	PT08
	<chr>	<chr>	<dbl>	<int>	<int>	<dbl>	<int>	<int>	<int>
1	10/03/2004	18.00.00	2.6	1360	150	11.9	1046	166	1056
2	10/03/2004	19.00.00	2.0	1292	112	9.4	955	103	1174
3	10/03/2004	20.00.00	2.2	1402	88	9.0	939	131	1140
4	10/03/2004	21.00.00	2.2	1376	80	9.2	948	172	1092
5	10/03/2004	22.00.00	1.6	1272	51	6.5	836	131	1205
6	10/03/2004	23.00.00	1.2	1197	38	4.7	750	89	1337

A data.frame: 6 × 15

	CO.GT.	PT08.S1.CO.	NMHC.GT.	C6H6.GT.	PT08.S2.NMHC.	NOx.GT.	PT08.S3.NOx.	NO2.GT.	PT0
	<dbl>	<int>	<int>	<dbl>	<int>	<int>	<int>	<int>	<in
1	2.6	1360	150	11.9	1046	166	1056	113	169
2	2.0	1292	112	9.4	955	103	1174	92	155
3	2.2	1402	88	9.0	939	131	1140	114	155
4	2.2	1376	80	9.2	948	172	1092	122	158
5	1.6	1272	51	6.5	836	131	1205	116	149
6	1.2	1197	38	4.7	750	89	1337	96	139

A data.frame: 6 × 13

2. Mendefinisikan variabel dependen dan variabel independen

```
head(Y)
```

	CO.GT.	PT08.S1.CO.	NMHC.GT.	C6H6.GT.	PT08.S2.NMHC.	NOx.GT.	PT08.S3.NOx.	NO2.GT.	PT0
	<dbl>	<int>	<int>	<dbl>	<int>	<int>	<int>	<int>	<in
1	2.6	1360	150	11.9	1046	166	1056	113	169
2	2.0	1292	112	9.4	955	103	1174	92	155
3	2.2	1402	88	9.0	939	131	1140	114	155
4	2.2	1376	80	9.2	948	172	1092	122	158
5	1.6	1272	51	6.5	836	131	1205	116	149
6	1.2	1197	38	4.7	750	89	1337	96	139

A data.frame: 6 × 10

	T	RH	AH
	<dbl>	<dbl>	<dbl>
1	13.6	48.9	0.7578
2	13.3	47.7	0.7255
3	11.9	54.0	0.7502
4	11.0	60.0	0.7867
5	11.2	59.6	0.7888
6	11.2	59.2	0.7848

A data.frame: 6 × 3

3. Mencari korelasi kanonik

$Xcor

	CO.GT.	PT08.S1.CO.	NMHC.GT.	C6H6.GT.	PT08.S2.NMHC.	NOx.GT.
CO.GT.	1.00000000	0.04141141	0.128351167	-0.031378275	0.02992582	0.526450925
PT08.S1.CO.	0.04141141	1.00000000	0.170007292	0.852687377	0.93310174	0.277992549
NMHC.GT.	0.12835117	0.17000729	1.000000000	0.037322518	0.11010353	-0.004427252
C6H6.GT.	-0.03137828	0.85268738	0.037322518	1.000000000	0.76743306	-0.001173976
PT08.S2.NMHC.	0.02992582	0.93310174	0.110103528	0.767433062	1.00000000	0.331272285
NOx.GT.	0.52645093	0.27799255	-0.004427252	-0.001173976	0.33127228	1.000000000
PT08.S3.NOx.	-0.08998085	0.08701943	0.048820888	0.512192702	-0.07366737	-0.436084134
NO2.GT.	0.67112706	0.15402953	0.103307496	-0.010992484	0.17648777	0.817139096
PT08.S4.NO2.	-0.07372396	0.84514865	0.162679943	0.774673427	0.87478154	0.035545972
PT08.S5.O3.	0.08030965	0.89243443	0.101185073	0.641334339	0.90990501	0.461888867

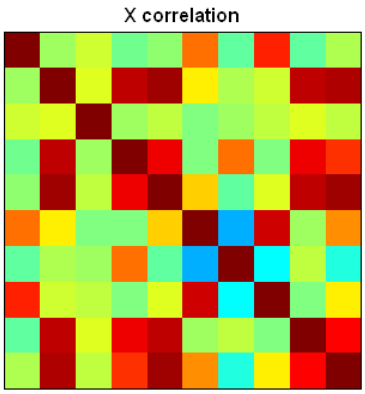
A matrix: 10 × 10 of type dbl

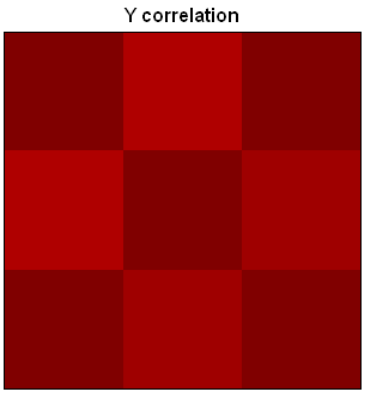
$Ycor

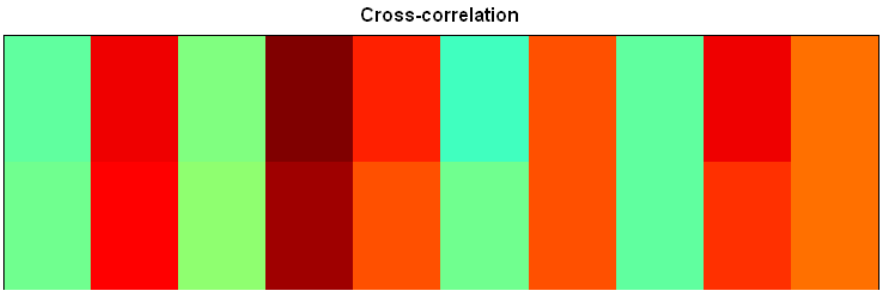
	T	RH	AH
T	1.0000000	0.8859105	0.9845466
RH	0.8859105	1.0000000	0.9368152
AH	0.9845466	0.9368152	1.0000000

A matrix: 3 × 3 of type dbl

$XYcor







[32]

```
cc1
```

$cor
0.997524513706709 · 0.652071979135027 · 0.219042271633645

$xcoef

CO.GT.	-5.181593e-07	-2.710346e-06	-1.180104e-05	9.307639e-05	6.269183e-05	-1.360517e
PT08.S1.CO.	6.879936e-07	-3.742827e-05	7.036851e-05	-4.542241e-05	-5.354287e-05	-4.740064e
NMHC.GT.	1.478453e-06	-8.361412e-06	2.064494e-05	2.085088e-05	1.478504e-05	2.256136e-
C6H6.GT.	-3.115955e-04	-1.434718e-04	-7.114564e-04	2.941522e-04	1.078658e-04	2.887281e-
PT08.S2.NMHC.	1.397350e-05	8.680649e-05	5.461489e-05	1.107082e-06	3.115059e-05	1.146963e-
NOx.GT.	-4.861564e-07	-4.001727e-05	-1.930712e-05	3.263337e-06	2.737813e-05	4.627291e-
PT08.S3.NOx.	3.210577e-06	4.232692e-06	4.661990e-05	3.254384e-06	-4.106805e-06	3.182382e-
NO2.GT.	1.127579e-06	6.563628e-05	3.798284e-05	-4.912486e-06	-4.512199e-05	2.274383e-
PT08.S4.NO2.	-4.975992e-06	1.350935e-07	-2.126695e-05	-3.798022e-06	-8.377193e-07	1.074667e-
PT08.S5.O3.	4.274189e-07	-2.334127e-05	-1.330351e-05	2.248257e-06	1.635167e-05	-8.289329e

A matrix: 10 × 10 of type dbl

$ycoef

T	-1.449412e-04	4.444639e-04	0.0016339732
RH	-5.753540e-05	-4.452916e-04	0.0005598961
AH	-3.375124e-05	6.402148e-05	-0.0024009247

A matrix: 3 × 3 of type dbl

$xcenter
CO.GT.: -34.207523778989 **PT08.S1.CO.:** 1048.99006091696 **NMHC.GT.:** -159.090092978519 **C6H6.GT.:** 1.86568344554879

[33]

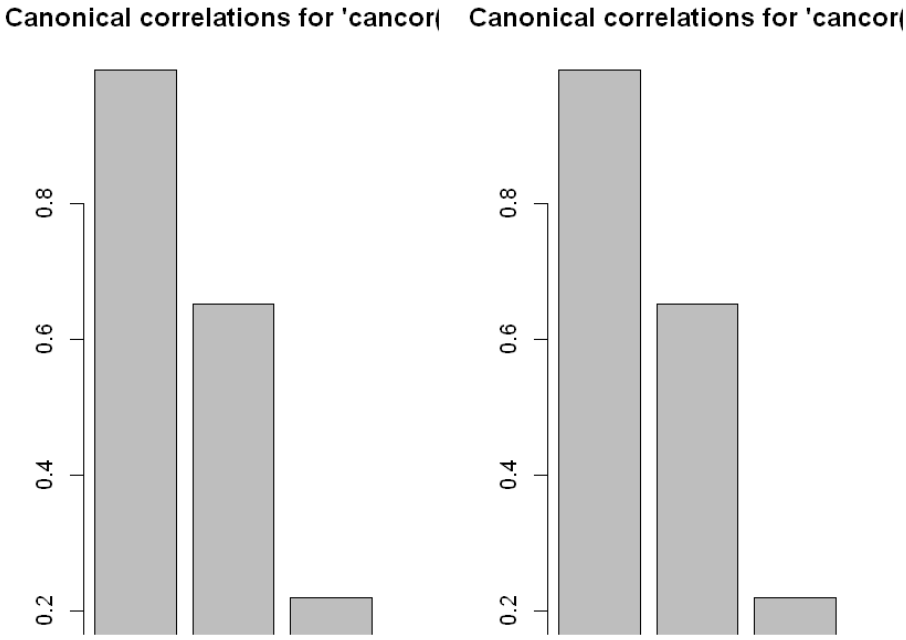
```
cc2
```

$cor
0.997524513706714 · 0.652071979135029 · 0.21904227163364

$names
$Xnames
'CO.GT.' · 'PT08.S1.CO.' · 'NMHC.GT.' · 'C6H6.GT.' · 'PT08.S2.NMHC.' · 'NOx.GT.' · 'PT08.S3.NOx.' · 'NO2.GT.' · 'PT08.S4.NO2.' · 'PT08.S5.O3.'

$Ynames
'T' · 'RH' · 'AH'

$ind.names
'1' · '2' · '3' · '4' · '5' · '6' · '7' · '8' · '9' · '10' · '11' · '12' · '13' · '14' · '15' · '16' · '17' · '18' · '19' · '20' · '21' · '22' · '23' · '24' · '25' · '26' · '27' · '28' · '29' · '30' · '31' · '32' · '33' · '34' · '35' · '36' · '37' · '38' · '39' · '40' · '41' · '42' · '43' · '44' · '45' · '46' · '47' · '48' · '49' · '50' · '51' · '52' · '53' · '54' · '55' · '56' · '57' · '58' · '59' · '60' · '61' · '62' · '63' · '64' · '65' · '66' · '67' · '68' · '69' · '70' · '71' · '72' · '73' · '74' · '75' · '76' · '77' · '78' · '79' · '80' · '81' · '82' · '83' · '84' · '85' · '86' · '87' · '88' · '89' · '90' · '91' · '92' · '93' · '94' · '95' · '96' · '97' · '98' · '99' · '100' · '101' · '102' · '103' · '104' · '105' · '106' · '107' · '108' · '109' · '110' · '111' · '112' · '113' · '114' · '115' · '116' · '117' · '118' · '119' · '120' · '121' · '122' · '123' · '124' · '125' · '126' · '127' · '128' · '129' · '130' · '131' · '132' · '133' · '134' · '135' · '136' · '137' · '138' · '139' · '140' · '141' · '142' · '143' · '144' · '145' · '146' · '147' · '148' · '149' · '150' · '151' · '152' · '153' · '154' · '155' · '156' · '157' · '158' · '159' · '160' · '161' · '162' · '163' · '164' · '165' · '166' · '167' · '168' · '169' · '170' · '171' · '172' · '173' · '174' · '175' · '176' · '177' · '178' · '179' · '180' · '181' · '182' · '183' · '184' · '185' · '186' · '187' · '188' · '189' · '190' · '191' · '192' · '193' · '194' · '195' · '196' · '197' · '198' · '199' · '200' · ... · '9158' · '9159' · '9160' · '9161' · '9162' · '9163' · '9164' · '9165' · '9166' · '9167' · '9168' · '9169' · '9170' · '9171' · '9172' · '9173' · '9174' · '9175' · '9176' · '9177' · '9178' · '9179' · '9180' · '9181' · '9182' · '9183' · '9184' · '9185' · '9186' · '9187' · '9188' · '9189' · '9190' · '9191' · '9192' · '9193' · '9194' · '9195' · '9196' · '9197' · '9198' · '9199' · '9200' · '9201' · '9202' · '9203' · '9204' · '9205' · '9206' · '9207' · '9208' · '9209' · '9210' · '9211' · '9212' · '9213' · '9214' · '9215' · '9216' · '9217' · '9218' · '9219' · '9220' · '9221' · '9222' · '9223' · '9224' · '9225' · '9226' · '9227' · '9228' · '9229' · '9230' · '9231' · '9232' · '9233' · '9234' · '9235' · '9236' · '9237' · '9238' · '9239' · '9240' · '9241' · '9242' · '9243' · '9244' · '9245' · '9246' · '9247' · '9248' · '9249' · '9250' · '9251' · '9252' · '9253' · '9254' · '9255' · '9256' · '9257' · '9258' · '9259' · '9260' · '9261' · '9262' · '9263' · '9264' · '9265' · '9266' · '9267' · '9268' · '9269' · '9270' · '9271' · '9272' · '9273' · '9274' · '9275' · '9276' · '9277' · '9278' · '9279' · '9280' · '9281' · '9282' · '9283' · '9284' · '9285' · '9286' · '9287' · '9288' · '9289' · '9290' · '9291' · '9292' · '9293' · '9294' · '9295' · '9296' · '9297' · '9298' · '9299' · '9300' · '9301' · '9302' · '9303' · '9304' · '9305' · '9306' · '9307' · '9308' · '9309' · '9310' · '9311' · '9312' · '9313' · '9314' · '9315' · '9316' · '9317' · '9318' · '9319' · '9320' · '9321' · '9322' · '9323' · '9324' ·

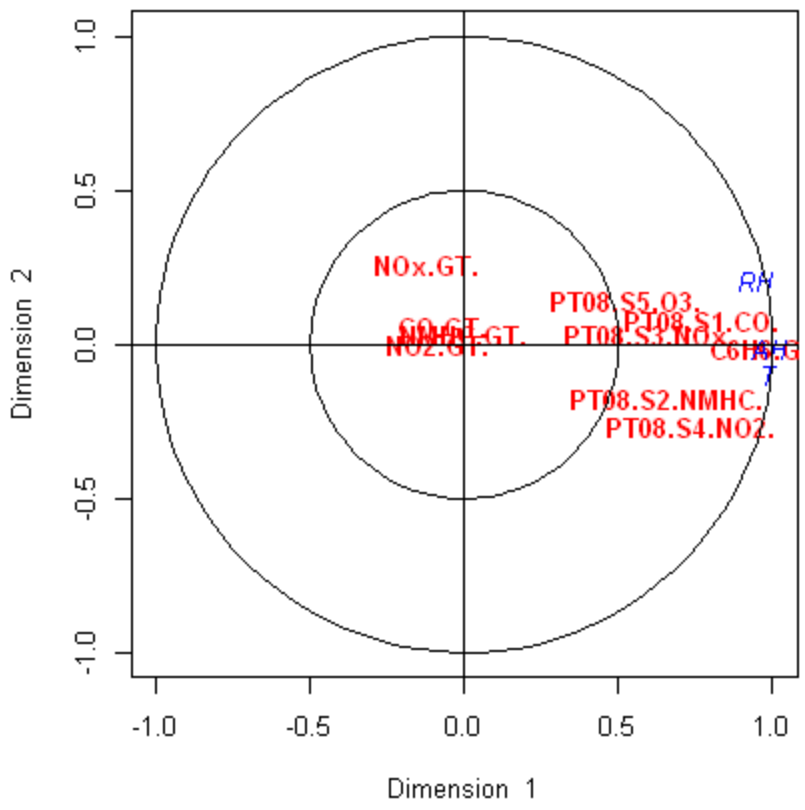


[11]

```
cc1$xcoef  ### function from standard R instalation
```

CO.GT.	-5.181593e-07	-2.710346e-06	-1.180104e-05	9.307639e-05	6.269183e-05	-1.360517e
PT08.S1.CO.	6.879936e-07	-3.742827e-05	7.036851e-05	-4.542241e-05	-5.354287e-05	-4.740064e
NMHC.GT.	1.478453e-06	-8.361412e-06	2.064494e-05	2.085088e-05	1.478504e-05	2.256136e-
C6H6.GT.	-3.115955e-04	-1.434718e-04	-7.114564e-04	2.941522e-04	1.078658e-04	2.887281e-
PT08.S2.NMHC.	1.397350e-05	8.680649e-05	5.461489e-05	1.107082e-06	3.115059e-05	1.146963e-
NOx.GT.	-4.861564e-07	-4.001727e-05	-1.930712e-05	3.263337e-06	2.737813e-05	4.627291e-
PT08.S3.NOx.	3.210577e-06	4.232692e-06	4.661990e-05	3.254384e-06	-4.106805e-06	3.182382e-
NO2.GT.	1.127579e-06	6.563628e-05	3.798284e-05	-4.912486e-06	-4.512199e-05	2.274383e-
PT08.S4.NO2.	-4.975992e-06	1.350935e-07	-2.126695e-05	-3.798022e-06	-8.377193e-07	1.074667e-
PT08.S5.O3.	4.274189e-07	-2.334127e-05	-1.330351e-05	2.248257e-06	1.635167e-05	-8.289329e

A matrix: 10 × 10 of type dbl




```
ccan <- candisc::cancor(X,Y)
summary(ccan)
```

```
Canonical correlation analysis of:
        10   X  variables:  CO.GT., PT08.S1.CO., NMHC.GT., C6H6.GT., PT08.S2.NMHC., NOx.GT., PT08.S3.NOx., NO2.GT., PT08.S4.NO2.,
  with   3   Y  variables:  T, RH, AH

    CanR  CanRSQ    Eigen  percent    cum                          scree
1 0.9975 0.99506 201.2308 99.60889  99.61 ******************************
2 0.6521 0.42520   0.7397  0.36616  99.98
3 0.2190 0.04798   0.0504  0.02495 100.00

Test of H0: The canonical correlations in the
current row and all that follow are zero
```

[28]

```
# ANALISIS KORELASI KANONIK
ccan <- candisc::cancor(X,Y)
summary(ccan)
```

```
Test of H0: The canonical correlations in the
current row and all that follow are zero

     CanR LR test stat approx F numDF denDF   Pr(> F)
1 0.99752      0.00271   5938.3    30 27427 < 2.2e-16 ***
2 0.65207      0.54722    365.3    18 18690 < 2.2e-16 ***
3 0.21904      0.95202     58.9     8  9346 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Raw canonical coefficients

   X  variables:
                   Xcan1       Xcan2      Xcan3
CO.GT.          5.0120e-05  2.6216e-04  0.0011415
PT08.S1.CO.    -6.6547e-05  3.6203e-03 -0.0068065
NMHC.GT.       -1.4301e-04  8.0877e-04 -0.0019969
C6H6.GT.        3.0140e-02  1.3878e-02  0.0688166
PT08.S2.NMHC. -1.3516e-03 -8.3965e-03 -0.0052827
NOx.GT.         4.7024e-05  3.8707e-03  0.0018675
PT08.S3.NOx.   -3.1055e-04 -4.0941e-04 -0.0045094
NO2.GT.        -1.0907e-04 -6.3488e-03 -0.0036739
PT08.S4.NO2.    4.8131e-04 -1.3067e-05  0.0020571
PT08.S5.O3.    -4.1343e-05  2.2577e-03  0.0012868

   Y  variables:
      Ycan1      Ycan2     Ycan3
T  0.0140196 -0.0429914 -0.158048
RH 0.0055652  0.0430715 -0.054157
AH 0.0032646 -0.0061926  0.232233
```

[31]

```
# Korelasi kanonik
res.cc <- cc(X,Y)
res.cc
```

$cor

0.997524513706714 · 0.652071979135029 · 0.21904227163364

$names

$Xnames

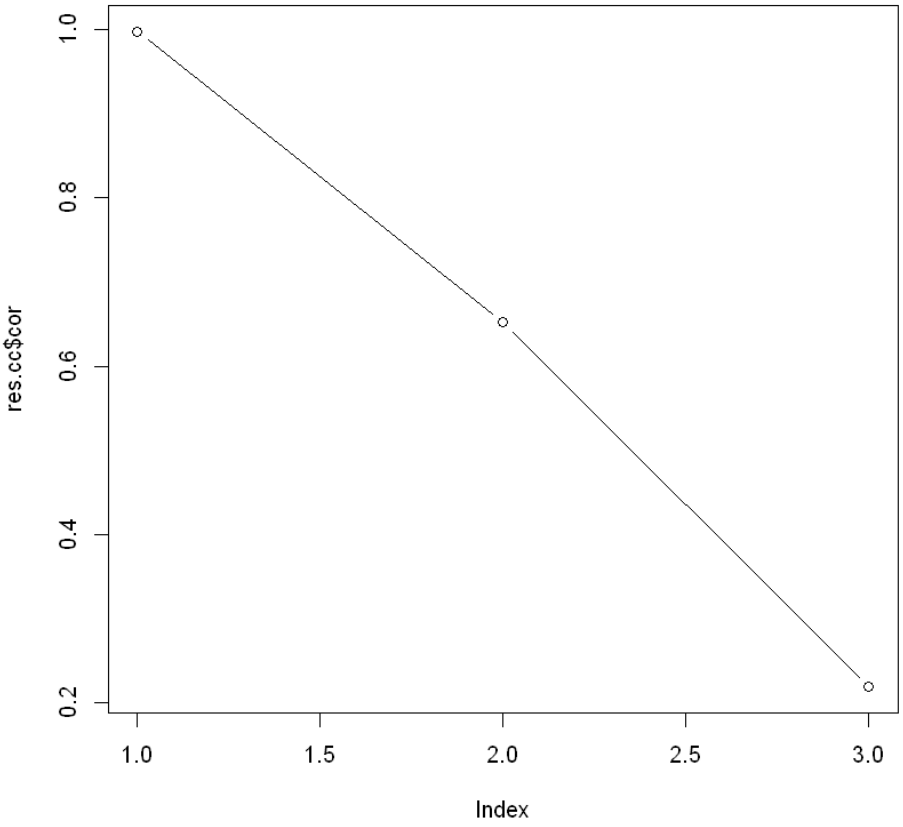
'CO.GT.' · 'PT08.S1.CO.' · 'NMHC.GT.' · 'C6H6.GT.' · 'PT08.S2.NMHC.' · 'NOx.GT.' · 'PT08.S3.NOx.' · 'NO2.GT.' · 'PT08.S4.NO2.' · 'PT08.S5.O3.'

$Ynames

'T' · 'RH' · 'AH'

$ind.names

'1' · '2' · '3' · '4' · '5' · '6' · '7' · '8' · '9' · '10' · '11' · '12' · '13' · '14' · '15' · '16' · '17' · '18' · '19' · '20' · '21' · '22' · '23' · '24' · '25' · '26' · '27' · '28' · '29' · '30' · '31' · '32' · '33' · '34' · '35' · '36' · '37' · '38' · '39' · '40' · '41' · '42' · '43' · '44' · '45' · '46' · '47' · '48' · '49' · '50' · '51' · '52' · '53' · '54' · '55' · '56' · '57' · '58' · '59' · '60' · '61' · '62' · '63' · '64' · '65' · '66' · '67' · '68' · '69' · '70' · '71' · '72' · '73' · '74' · '75' · '76' · '77' · '78' · '79' · '80' · '81' · '82' · '83' · '84' · '85' · '86' · '87' · '88' · '89' · '90' · '91' · '92' · '93' · '94' · '95' · '96' · '97' · '98' · '99' · '100' · '101' · '102' · '103' · '104' · '105' · '106' · '107' · '108' · '109' · '110' · '111' · '112' · '113' · '114' · '115' · '116' · '117' · '118' · '119' · '120' · '121' · '122' · '123' · '124' · '125' · '126' · '127' · '128' · '129' · '130' · '131' · '132' · '133' · '134' · '135' · '136' · '137' · '138' · '139' · '140' · '141' · '142' · '143' · '144' · '145' · '146' · '147' · '148' · '149' · '150' · '151' · '152' · '153' · '154' · '155' · '156' · '157' · '158' · '159' · '160' · '161' · '162' · '163' · '164' · '165' · '166' · '167' · '168' · '169' · '170' · '171' · '172' · '173' · '174' · '175' · '176' · '177' · '178' · '179' · '180' · '181' · '182' · '183' · '184' · '185' · '186' · '187' · '188' · '189' · '190' · '191' · '192' · '193' · '194' · '195' · '196' · '197' · '198' · '199' · '200' · ... · '9158' · '9159' · '9160' · '9161' · '9162' · '9163' · '9164' · '9165' · '9166' · '9167' · '9168' · '9169' · '9170' · '9171' · '9172' · '9173' · '9174' · '9175' · '9176' · '9177' · '9178' · '9179' · '9180' · '9181' · '9182' · '9183' · '9184' · '9185' · '9186' · '9187' · '9188' · '9189' · '9190' · '9191' · '9192' · '9193' · '9194' · '9195' · '9196' · '9197' · '9198' · '9199' · '9200' · '9201' · '9202' · '9203' · '9204' · '9205' · '9206' · '9207' · '9208' · '9209' · '9210' · '9211' · '9212' · '9213' · '9214' · '9215' · '9216' · '9217' · '9218' · '9219' · '9220' · '9221' · '9222' · '9223' · '9224' · '9225' · '9226' · '9227' · '9228' · '9229' · '9230' · '9231' · '9232' · '9233' · '9234' · '9235' · '9236' · '9237' · '9238' · '9239' · '9240' · '9241' · '9242' · '9243' · '9244' · '9245' · '9246' · '9247' · '9248' · '9249' · '9250' · '9251' · '9252' · '9253' · '9254' · '9255' · '9256' · '9257' · '9258' · '9259' · '9260' · '9261' · '9262' · '9263' · '9264' · '9265' · '9266' · '9267' · '9268' · '9269' · '9270' · '9271' · '9272' · '9273' · '9274' · '9275' · '9276' · '9277' · '9278' · '9279' · '9280' · '9281' · '9282' · '9283' · '9284' · '9285' · '9286' · '9287' · '9288' · '9289' · '9290' · '9291' · '9292' · '9293' · '9294' · '9295' · '9296' · '9297' · '9298' · '9299' · '9300' · '9301' · '9302' · '9303' · '9304' · '9305' · '9306' · '9307' · '9308' · '9309' · '9310' · '9311' · '9312' · '9313' · '9314' · '9315' · '9316' · '9317' · '9318' · '9319' · '9320' · '9321' · '9322' · '9323' · '9324' ·



$xcoef

CO.GT.	5.011969e-05	2.621620e-04	0.001141472
PT08.S1.CO.	-6.654716e-05	3.620302e-03	-0.006806494
NMHC.GT.	-1.430055e-04	8.087694e-04	-0.001996911
C6H6.GT.	3.013952e-02	1.387751e-02	0.068816630
PT08.S2.NMHC.	-1.351607e-03	-8.396481e-03	-0.005282703
NOx.GT.	4.702417e-05	3.870728e-03	0.001867508
PT08.S3.NOx.	-3.105477e-04	-4.094131e-04	-0.004509376
NO2.GT.	-1.090667e-04	-6.348762e-03	-0.003673944
PT08.S4.NO2.	4.813099e-04	-1.306711e-05	0.002057076
PT08.S5.O3.	-4.134270e-05	2.257718e-03	0.001286801

$ycoef

T	0.014019647	-0.042991401	-0.15804838
RH	0.005565193	0.043071464	-0.05415675
AH	0.003264637	-0.006192568	0.23223285

$corr.X.xscores

CO.GT.	-0.0645804385	0.059157894	-0.13650868
PT08.S1.CO.	0.7722444070	0.078624736	-0.27563771
NMHC.GT.	0.0006854782	0.032915112	-0.64762120
C6H6.GT.	0.9835130409	-0.012818726	-0.08521165
PT08.S2.NMHC.	0.6603197628	-0.175497135	-0.18629424
NOx.GT.	-0.1150420394	0.260970217	-0.14001608
PT08.S3.NOx.	0.5994652418	0.037386679	-0.17219057
NO2.GT.	-0.0863096466	-0.000638997	-0.23938468
PT08.S4.NO2.	0.7389719352	-0.265938442	-0.08947749
PT08.S5.O3.	0.5243624702	0.142343466	-0.15321747

$corr.Y.xscores

T	0.9860420	-0.097960842	-0.003921233
RH	0.9432468	0.212141692	-0.001011434
AH	0.9932182	-0.008550337	0.020127497

$corr.X.yscores

CO.GT.	-0.0644205705	0.0385752051	-0.02990117
PT08.S1.CO.	0.7703327266	0.0512689874	-0.06037631
NMHC.GT.	0.0006837813	0.0214630220	-0.14185642
C6H6.GT.	0.9810783679	-0.0083587317	-0.01866495
PT08.S2.NMHC.	0.6586851502	-0.1144367639	-0.04080631
NOx.GT.	-0.1147572544	0.1701713656	-0.03066944
PT08.S3.NOx.	0.5979812738	0.0243788061	-0.03771701
NO2.GT.	-0.0860959882	-0.0004166721	-0.05243536
PT08.S4.NO2.	0.7371426203	-0.1734110063	-0.01959935
PT08.S5.O3.	0.5230644181	0.0928181859	-0.03356110

$corr.Y.yscores

T	0.9884890	-0.15023011	-0.01790172
RH	0.9455876	0.32533478	-0.00461753
AH	0.9956830	-0.01311257	0.09188864

VI. Pembahasan

7.1 Analisis Klasifikasi

Dalam melakukan analisis klasifikasi, data yang kami miliki yaitu data red-wine harus kami tentukan variabel apa yang akan kami jadikan acuan dalam mengklasifikasi. Pada tugas ini kami memilih variabel "quality" yang akan kami gunakan. Sebelum memilih suatu metode kami harus melakukan uji normal untuk melihat apakah data yang kami miliki berdistribusi normal atau tidak. Hal ini bertujuan untuk mempermudah kami memilih metode analisis klasifikasi, karena terdapat perbedaan metode bila data yang kami miliki tidak berdistribusi normal. Dari uji normal menggunakan uji shapiro wilk, kami mengetahui bahwa data wine yang kami miliki tidak berdistribusi normal, sehingga metode yang kami pilih adalah metode decision tree dan KNN.

Penggunaan metode decision tree diawali dengan, kami memilih untuk mengklasifikasikan data wine yang kami miliki berdasarkan variabel "quality" di mana kami mengkategorikan data "quality" yang bernilai ≤ 5 akan dikategorikan sebagai "kurang baik" dan untuk data yang bernilai > 5 akan dikategorikan sebagai "cukup baik". Lalu kami membuat bagan pohon atau decision tree dengan menggunakan package tree pada R. Dengan decision tree, kita dapat melihat variabel apa saja dan berapa range nilai dari variabel tersebut untuk bisa dikatakan wine dengan nilai variabel tersebut dapat dikatakan memiliki kualitas yang cukup baik atau kurang baik. Setelah itu kami dapat memprediksi kualitas wine yang akan datang dengan data kualitas wine yang kami miliki sekarang, seperti yang dapat dilihat pada akhir bagian decision tree 6.1.

Penggunaan metode K-Nearest Neighbour diawali dengan melihat banyaknya data wine yang masuk ke dalam kategori tertentu, contohnya pada data yang kami miliki data yang dikategorikan cukup baik sebanyak 855, sedangkan data yang dikategorikan kurang baik sebanyak 744. Setelah mengetahui sebaran kategori pada data, kami memeriksa akurasi cluster dengan mencoba K=1 sebagai langkah awal, lalu kami mencoba K=7, dan dapat dilihat bahwa akurasi saat K=7 adalah 0,82 sedangkan saat K=1 tentunya hasilnya adalah 1 karena merupakan keseluruhan data. Kemudian kami tertarik untuk melihat bagaimana data tersebar di ketujuh kelompok sehingga dengan begitu kami menggunakan cara yang dinamakan dengan votes yang mengeluarkan hasil seperti pada akhir bagian KNN 6.1.

7.2 Korelasi Kanonik

Dalam melakukan analisis korelasi kanonik, yang pertama kami lakukan adalah menentukan variabel atau atribut mana dari data “Air Quality” yang akan kami punya untuk dijadikan variabel dependen dan variabel independen dalam proses analisis data. Pada kesempatan ini kami memilih T, RH, dan AH sebagai variabel dependen, sedangkan atribut yang lain adalah variabel independen karena mereka merupakan pembentuk dari T, RH, dan AH pada udara. Setelah menentukan variabel X dan Y dari data, kami membuat tabel korelasi antar pembentuk variabel untuk kedua variabel X dan Y, lalu kami membuat visualisasinya untuk memperjelas dalam melihat seberapa besar korelasi antar variabelnya.

Berdasarkan hasil dari analisis korelasi kanonik di atas, Korelasi kanonik pertama (korelasi antara pasangan pertama dari kanonik) adalah sebesar 0.9975 sedangkan squared canonical correlation adalah 0.99506. Nilai ini merepresentasikan korelasi tertinggi yang mungkin terjadi antara beberapa kombinasi linear yang terbentuk. Kontribusi keragaman yang dijelaskan oleh fungsi kanonik terlihat fungsi kanonik pertama menjelaskan keragaman total sebesar 99.7%. Sedangkan yang kedua 0.37%, dan ketiga 0.024%. Berdasarkan proporsi keragaman tersebut dengan menggunakan kriteria batasan minimal kontribusi keragaman 70%, maka cukup mengambil fungsi kanonik yang pertama saja karena keragaman yang mampu diterangkan sudah cukup besar. Dilihat dari Uji rasio kemungkinan (uji parsial) di atas juga menunjukkan bahwa korelasi kanonik yang pertama berbeda nyata dengan nol pada taraf nyata $\alpha = 5\%$ ($Pr < \alpha$) sehingga dapat disimpulkan bahwa pasangan peubah kanonik pertama memiliki korelasi yang signifikan dan dapat digunakan untuk menjelaskan hubungan antar variabel yang ditentukan dengan jelas. Lalu jika dilihat dari grafik yang tebentuk, hal ini juga mendukung pernyataan sebelumnya.

Dari tabel koefisien X dan Y dapat diperhatikan untuk fungsi kanonik pertama saja, dapat dilihat peubah X yang memberikan kontribusi terbesar adalah C6H6.GT. diikuti oleh PT08S2.NMHC dan PT08.S4.NO2. Sedangkan untuk peubah Y yang memberikan kontribusi terbesar adalah Temperature diikuti dengan RH (Relative Humidity) dan AH (Absolute Humidity).

Dari tabel korelasi scores dapat diketahui peubah Y paling berhubungan dengan fungsi kanonik pertama adalah AH atau Absolute Humidity (Kelembaban Mutlak) sedangkan peubah X yang paling berhubungan dengan fungsi kanonik pertama adalah C6H6.GT. atau True hourly averaged Benzene concentration in microg/m^3 (Konsentrasi Benzena rata-rata per jam sebenarnya dalam mikrog/m^3).

Untuk korelasi silang antar peubah-peubah dependen terhadap fungsi kanonik peubah independen yang berhubungan paling erat dengan fungsi kanonik pertama adalah AH atau Absolute Humidity. Sedangkan korelasi silang antar peubah-peubah independen terhadap fungsi kanonik peubah dependen yang berhubungan paling erat dengan fungsi kanonik pertama adalah C6H6.GT.