

Pokok Bahasan X

Association Rule

Kode Pokok Bahasan: TIK.RPL03.001.010.01

Deskripsi Pokok Bahasan:

Membahas tentang Association Rule pada R dengan dataset yang diberikan.

No	Elemen Kompetensi	Indikator Kinerja	Jml Jam	Hal
1	Memahami cara implementasi Association Rule di R	1.1 Mampu memahami cara implementasi Association Rule data kecelakaan pesawat pada R dan excel 1.2 Menerapkan Association Rule pada data nasabah	1	12

TUGAS PENDAHULUAN

Hal yang harus dilakukan dan acuan yang harus dibaca sebelum praktikum :

1. Menginstal R pada PC masing-masing praktikan.
2. Menginstal R Studio pada PC masing-masing praktikan.

DAFTAR PERTANYAAN

1. Berikan penjelasan mengenai apa itu "Association Rule with R and Excel"?

Aturan asosiasi dibuat dengan menganalisis data untuk pola jika/maka yang sering dan menggunakan dukungan kriteria dan keyakinan untuk mengidentifikasi hubungan yang paling penting. Dukungan merupakan indikasi seberapa sering item muncul dalam database.

Bentuk aturan asosiasi adalah $\{X \rightarrow Y\}$, di mana $\{X\}$ adalah sekumpulan item dan $\{Y\}$ adalah item. Implikasi dari aturan asosiasi ini adalah bahwa jika semua item di $\{X\}$ muncul di beberapa keranjang, maka $\{Y\}$ "kemungkinan besar" juga akan muncul di keranjang itu. $\{Y\}$ disebut konsekuensi atau sisi kanan (RHS).

2. Apa kelebihan Association Rule dibanding algoritma lainnya?

Kelebihan asosiasi rule dengan apriori ini adalah lebih sederhana dan dapat menangani data yang besar. Sedangkan algoritma lainnya memiliki kelemahan dalam penggunaan memori saat jumlah data besar, tentunya berpengaruh terhadap banyaknya item yang diproses.



3. Sebutkan package yang dibutuhkan untuk melakukan Association Rule with R !

```
library(arules)
library(arulesViz)
```

TEORI SINGKAT

Algoritma apriori merupakan salah satu algoritma klasik data mining. Algoritma apriori digunakan agar komputer dapat mempelajari aturan asosiasi, mencari pola hubungan antar satu atau lebih item dalam suatu dataset.

Algoritma apriori banyak digunakan pada data transaksi atau biasa disebut market basket, misalnya sebuah swalayan memiliki market basket, dengan adanya algoritma apriori, pemilik swalayan dapat mengetahui pola pembelian seorang konsumen, jika seorang konsumen membeli item A , B, punya kemungkinan 50% dia akan membeli item C, pola ini sangat signifikan dengan adanya data transaksi selama ini.

Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, support (nilai penunjang) yaitu persentase kombinasi item tersebut dalam database dan confidence (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiatif.

LAB SETUP

Hal yang harus disiapkan dan dilakukan oleh praktikan untuk menjalankan praktikum modul ini.

1. Menginstall library yang dibutuhkan untuk mengerjakan modul.
2. Menjalankan R Studio.

ELEMEN KOMPETENSI I

Deskripsi:

Memahami cara implementasi Association Rule di R

Kompetensi Dasar:

- 1 Mampu memahami cara implementasi Association Rule data kecelakaan pesawat pada R
- 2 Menerapkan Association Rule pada data nasabah.

Latihan 1.1.1



Penjelasan Singkat :

Pada latihan ini anda akan diminta untuk menerapkan Association Rule pada data kecelakaan pesawat menggunakan R studio.

Langkah-Langkah Praktikum:

1. Berikut dibawah ini adalah langkah-langkah Association Rules dengan menggunakan R.

Berikut dibawah ini adalah data yang digunakan

perusahaan	sebab	hasil
A	Kerusakan Pesawat	Semua Selamat
A	Kerusakan Pesawat	Sebagian Selamat
B	Kerusakan Pesawat	Semua Selamat
B	Kerusakan Pesawat	Semua Selamat
A	Cuaca Buruk	Semua Selamat
A	Cuaca Buruk	Sebagian Selamat
A	Cuaca Buruk	Semua Selamat
A	Cuaca Buruk	Semua Selamat
C	Cuaca Buruk	Tidak Selamat
C	Tidak Jelas	Semua Selamat
A	Tidak Jelas	Semua Selamat
A	Tidak Jelas	Semua Selamat

2. Jalankan perintah pada script R berikut :

Script R

```
library(arules)
install.packages("arulesViz")
library(arulesViz)
dataku=read.delim("clipboard")
categoric<- c("perusahaan", "sebab", "hasil")

transactions <- as(dataku[, categoric], "transactions")
itemFrequencyPlot(transactions, topN = 7)

rules <- apriori(transactions,parameter = list(minlen=3, supp=0.10, conf=0.7), appearance
= list(rhs=c("hasil=Semua Selamat", "hasil=Sebagian Selamat", "hasil=Tidak Selamat")))

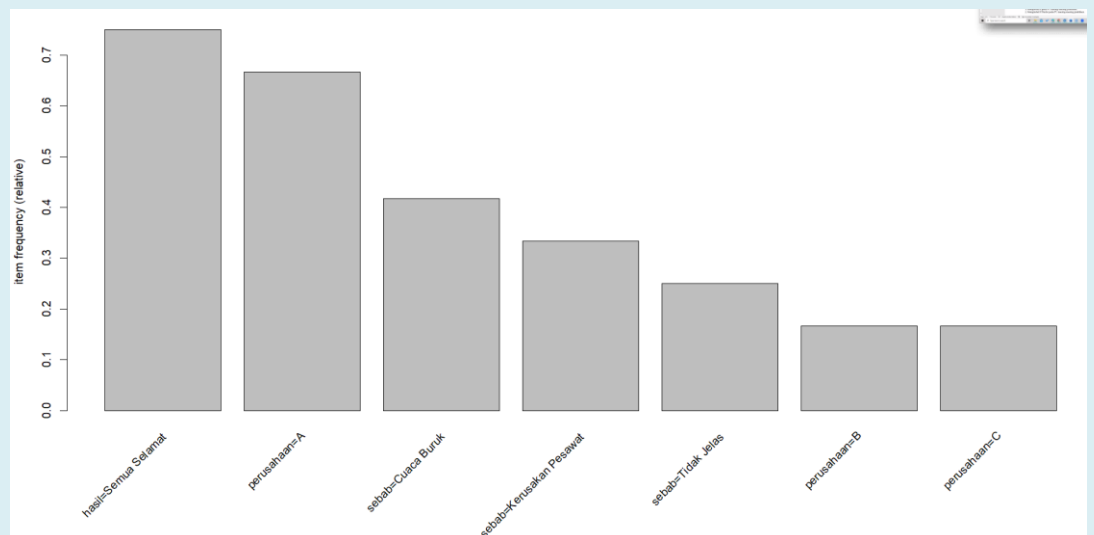
inspect(rules)
```

```
plot(rules, method="graph")
```

```
plot(rules, method="graph", engine = "interactive")
```

3. Tampilkan output pada kolom berikut :

Output :



```
> rules <- apriori(transactions,parameter = list(minlen=3, supp=0.10,
  conf=0.7), appearance = list(rhs=c("hasil=Semua Selamat","hasil=Sebag
  ian Selamat","hasil=Tidak Selamat")))
```

Apriori

Parameter specification:

confidence	minval	smax	arem	aval	originalSupport	maxtime
0.7	0.1	1	none	FALSE	TRUE	5

support	minlen	maxlen	target	ext
0.1	3	10	rules	TRUE

Algorithmic control:

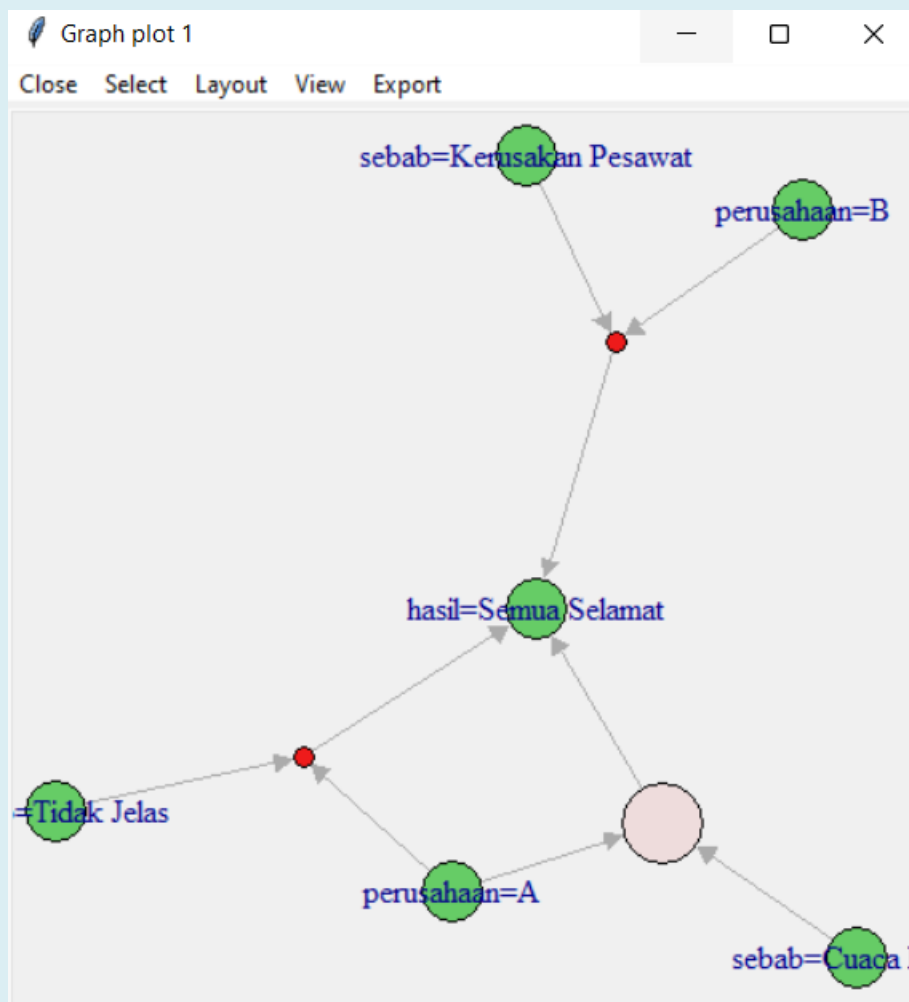
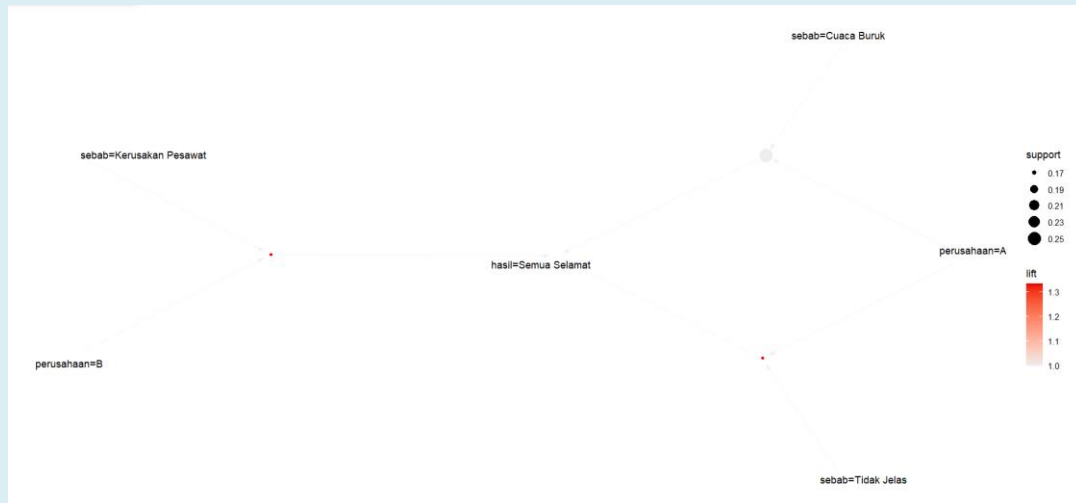
filter	tree	heap	memopt	load	sort	verbose
0.1	TRUE	TRUE	FALSE	TRUE	2	TRUE

Absolute minimum support count: 1

```
set item appearances ...[3 item(s)] done [0.00s].
set transactions ...[9 item(s), 12 transaction(s)] done [0.00s].
sorting and recoding items ... [8 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 done [0.00s].
writing ... [3 rule(s)] done [0.00s].
creating S4 object ... done [0.00s].
```

```
> |
```

```
> inspect(rules)
lhs      rhs      support confidence coverage lift count
[1] {perusahaan=B,      {hasil=Semua Selamat} 0.1666667      1.00 0.1666667 1.333333 2
    sebab=Kerusakan Pesawat}
[2] {perusahaan=A,      {hasil=Semua Selamat} 0.1666667      1.00 0.1666667 1.333333 2
    sebab=Tidak Jelas}
[3] {perusahaan=A,      {hasil=Semua Selamat} 0.2500000      0.75 0.3333333 1.000000 3
    sebab=Cuaca Buruk}
```



4. Jelaskan perbedaan hasilnya bila menggunakan Excel :

Perbedaan umum yang didapatkan bisa dilihat jika menggunakan excel kita harus melakukan perhitungan dengan cara manual yang lebih memakan waktu, jika menggunakan R maka akan lebih mudah mengolah data yang kita inginkan dengan metode-metode clustering yang ada.

5. Susun data secara bertumpuk (stacked) dengan dua kolom yaitu kejadian dan peristiwa. Kolom kejadian diberi nilai 1 untuk peristiwa yang sama. Lengkapi datanya, contoh :

Kejadian	Peristiwa
1	A
1	Kerusakan Pesawat
1	Semua Selamat
2	A
2	Kerusakan Pesawat
2	Sebagian Selamat
3	B
3	Kerusakan Pesawat
3	Semua Selamat
4	B
4	Kerusakan Pesawat
4	Semua Selamat
5	A
5	Cuaca Buruk
5	Semua Selamat
6	A
6	Cuaca Buruk
6	Sebagian Selamat
7	A
7	Cuaca Buruk
7	Semua Selamat
8	A
8	Cuaca Buruk
8	Semua Selamat
9	C
9	Cuaca Buruk
9	Tidak Selamat
10	C
10	Tidak Jelas
10	Semua Selamat

11	A
11	Tidak Jelas
11	Semua Selamat
12	A
12	Tidak Jelas
12	Semua Selamat

6. Lakukan dengan perintah sebagai berikut :

Script R :

```
#Install Packages
install.packages("arules")
install.packages("arulesViz")

#memanggil packages
library(arules)
library(arulesViz)

#memanggil data
latihan=read.delim("clipboard")
latihan

#analisis data
latihan1<-split(latihan$peristiwa,latihan$kejadian)
latihan2<-as(latihan1,"transactions")
latihan2

rules<-apriori(latihan2,parameter=list
(supp=0.04,conf=0.9,minlen=3))

inspect(rules)

itemFrequencyPlot(latihan2, topN = 7)

plot(rules, method="graph")
```

Output :

```
> #analisis data
> latihan1<-split(latihan$Peristiwa,latihan$Kejadian)
> latihan2<-as(latihan1,"transactions")
> latihan2
transactions in sparse format with
12 transactions (rows) and
9 items (columns)
> rules<-apriori(latihan2,parameter=list
+               (supp=0.04,conf=0.9,minlen=3))
Apriori

Parameter specification:
confidence minval smax arem aval originalSupport maxtime support minlen maxlen target ext
0.9 0.1 1 none FALSE TRUE 5 0.04 3 10 rules TRUE

Algorithmic control:
filter tree heap memopt load sort verbose
0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 0

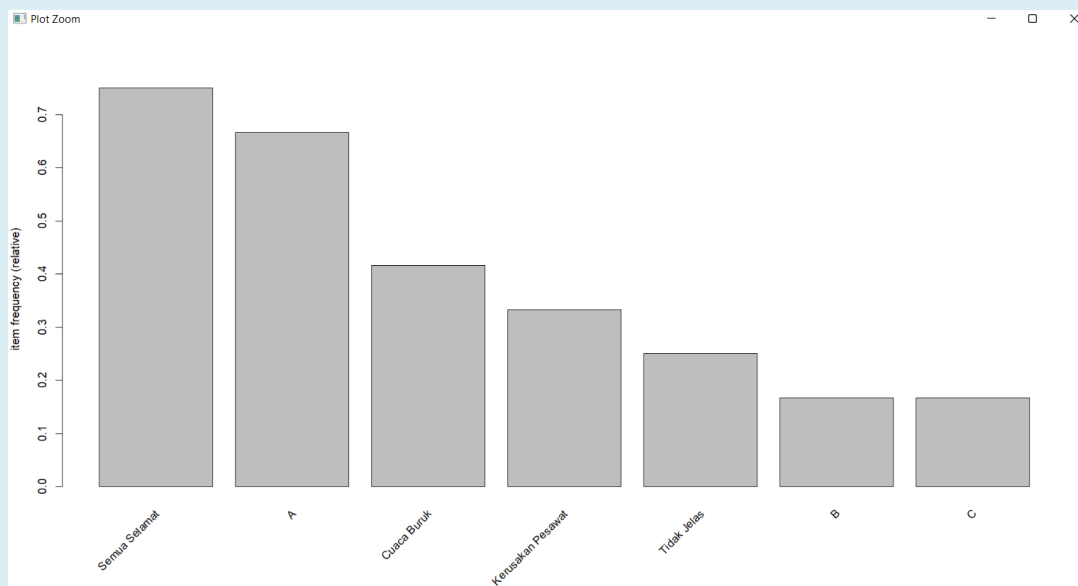
set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[9 item(s), 12 transaction(s)] done [0.00s].
sorting and recoding items ... [9 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 done [0.00s].
writing ... [11 rule(s)] done [0.00s].
creating S4 object ... done [0.00s].
> |
```

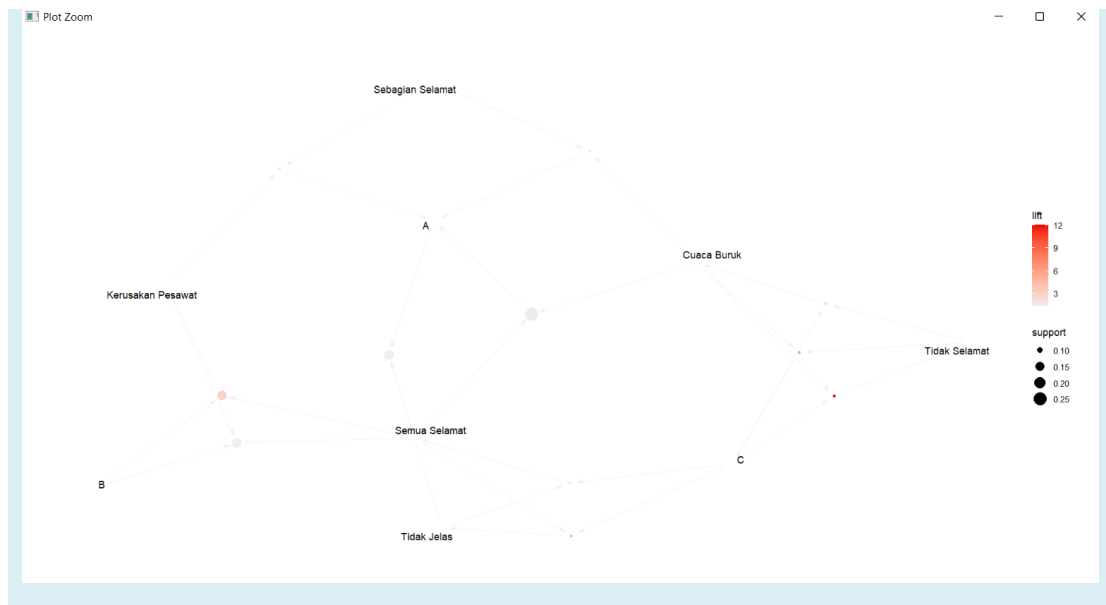
```
> inspect(rules)
```

	lhs	rhs	support	confidence	coverage
[1]	{C, Tidak Selamat}	=> {Cuaca Buruk}	0.08333333	1	0.08333333
[2]	{Cuaca Buruk, Tidak Selamat}	=> {C}	0.08333333	1	0.08333333
[3]	{C, Cuaca Buruk}	=> {Tidak Selamat}	0.08333333	1	0.08333333
[4]	{Kerusakan Pesawat, Sebagian Selamat}	=> {A}	0.08333333	1	0.08333333
[5]	{Cuaca Buruk, Sebagian Selamat}	=> {A}	0.08333333	1	0.08333333
[6]	{B, Kerusakan Pesawat}	=> {Semua Selamat}	0.16666667	1	0.16666667
[7]	{B, Semua Selamat}	=> {Kerusakan Pesawat}	0.16666667	1	0.16666667
[8]	{C, Tidak Jelas}	=> {Semua Selamat}	0.08333333	1	0.08333333
[9]	{C, Semua Selamat}	=> {Tidak Jelas}	0.08333333	1	0.08333333
[10]	{A, Tidak Jelas}	=> {Semua Selamat}	0.16666667	1	0.16666667
[11]	{Cuaca Buruk, Semua Selamat}	=> {A}	0.25000000	1	0.25000000

	lift	count
[1]	2.400000	1
[2]	6.000000	1
[3]	12.000000	1
[4]	1.500000	1
[5]	1.500000	1
[6]	1.333333	2
[7]	3.000000	2
[8]	1.333333	1
[9]	4.000000	1
[10]	1.333333	2
[11]	1.500000	3

```
> |
```





Latihan 1.1.2

Penjelasan Singkat :

Pada latihan ini anda akan diminta untuk menerapkan Association Rule pada data nasabah menggunakan R studio.

Tugas :

Berikut dibawah ini adalah data mengenai 11 nasabah pada koperasi XYZ dengan status kredit yang macet (Not Succeed) dan tidak macet (Succeed). Bangunlah association rule berdasarkan data dibawah menggunakan langkah-langkah di atas menggunakan R. Tetapkan nilai k-itemset dari support yang telah memenuhi minimum support dari k-itemset.

No	Graduate	Skill	Wellfare	Personalit y	Credit History
1	High School	Ok	Poor	Good	Succeed
2	Primary School	Not Ok	Very Poor	Good	Not Succeed
3	Primary School	Ok	Poor	Moderate	Not Succeed
4	High School	Ok	Poor	Good	Succeed
5	Primary School	Not Ok	Poor	Good	Not Succeed
6	Primary School	Ok	Very Poor	Moderate	Not Succeed
7	Primary School	Ok	Poor	Bad	Not Succeed
8	Primary School	Not Ok	Poor	Good	Not Succeed
9	Primary School	Not Ok	Poor	Good	Not Succeed

10	Primary School	Ok	Poor	Bad	Not Succeed
11	High School	Ok	Poor	Good	Succeed

Script R

```

dataibnu=read.delim("clipboard")
dataibnu
categoric<- c("Graduate", "Skill", "Wellfare", "Personality", "Credit.History")

transactions <- as(dataibnu[, categoric], "transactions")
itemFrequencyPlot(transactions, topN = 5)

rules <- apriori(transactions,parameter = list(minlen=3, supp=0.10, conf=0.7),
appearance      =      list(rhs=c("Credit.History=Succeed","Credit.History=Not
Succeed"))))

inspect(rules)

plot(rules, method="graph")

plot(rules, method="graph", engine = "interactive")

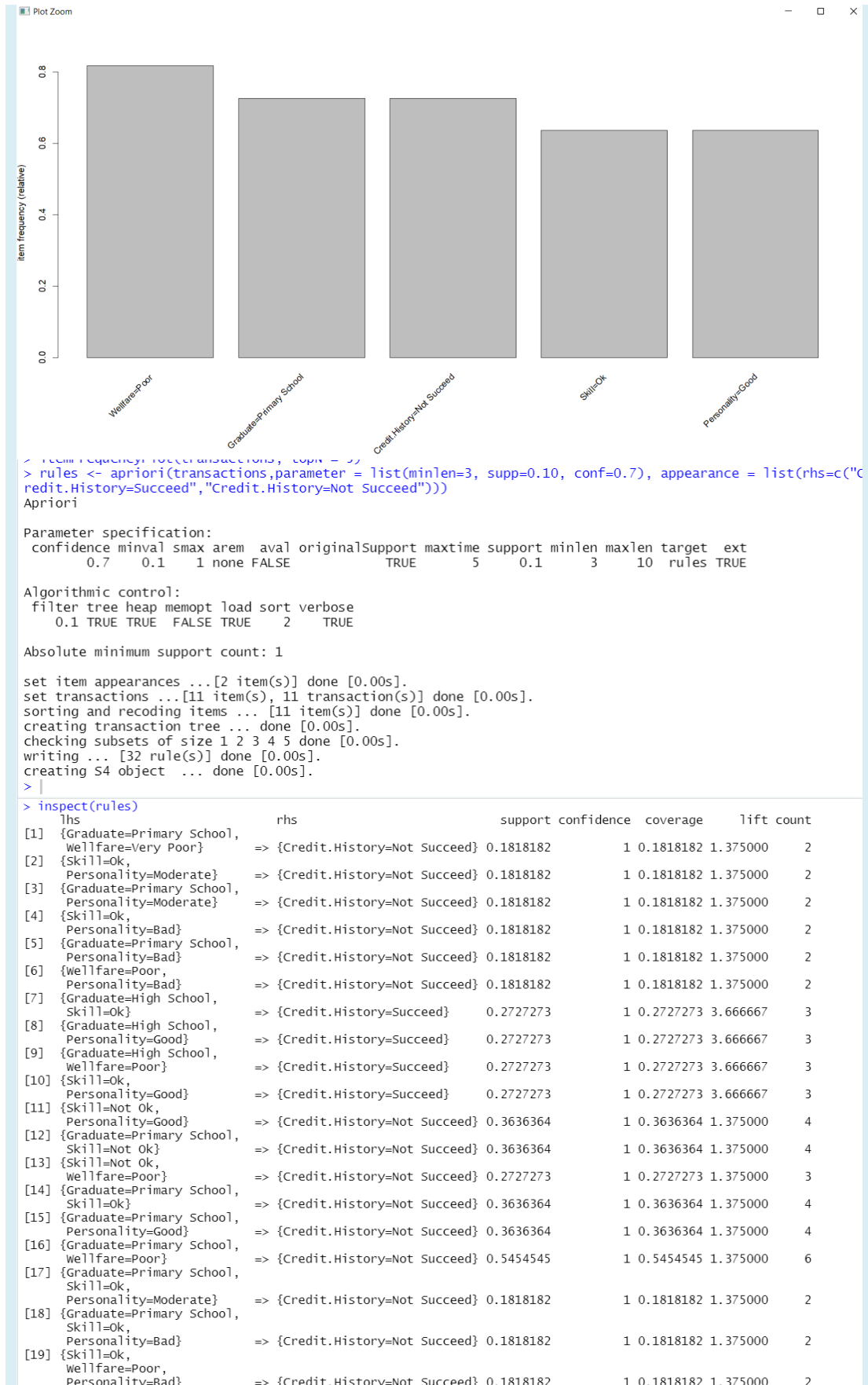
```

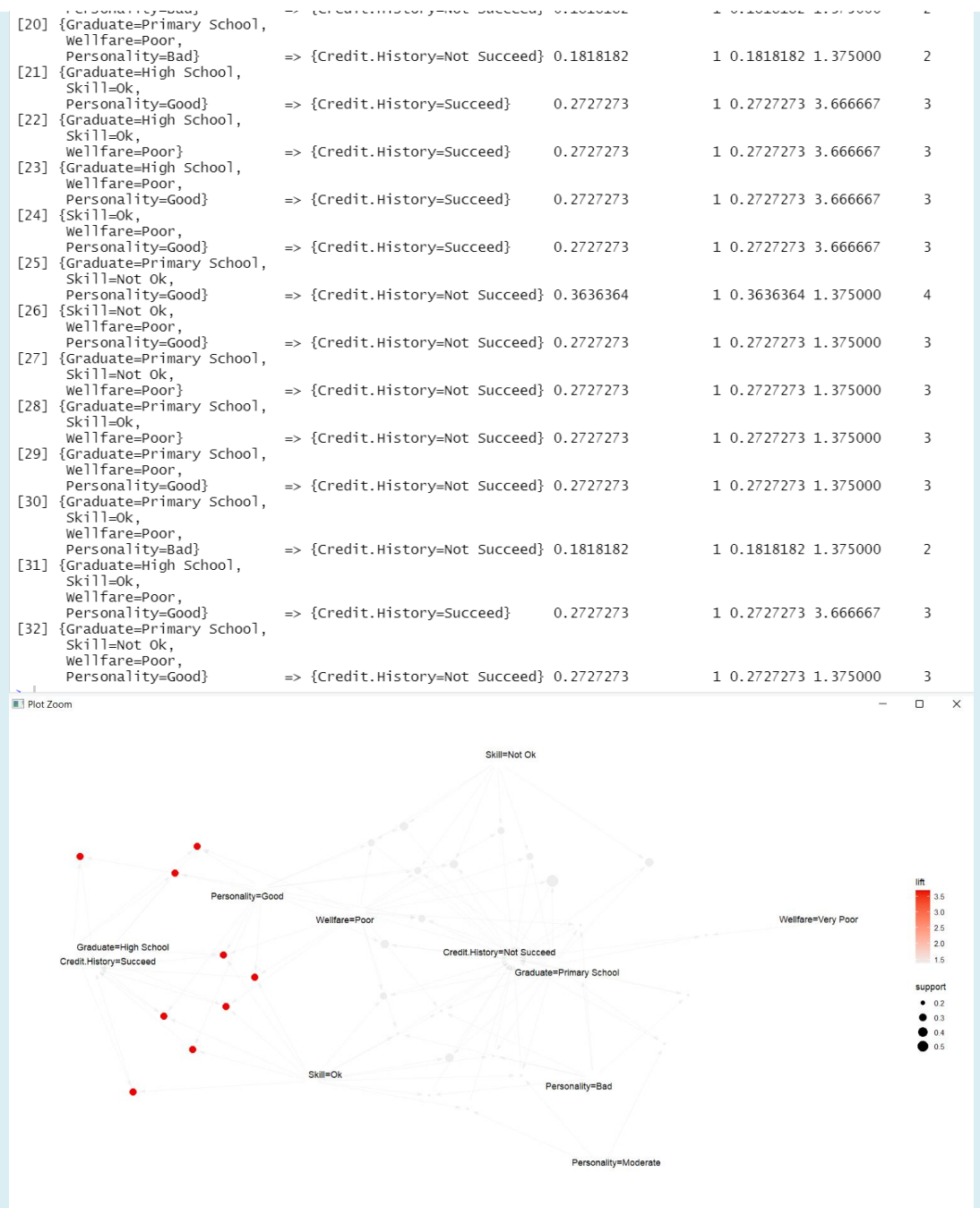
Output :

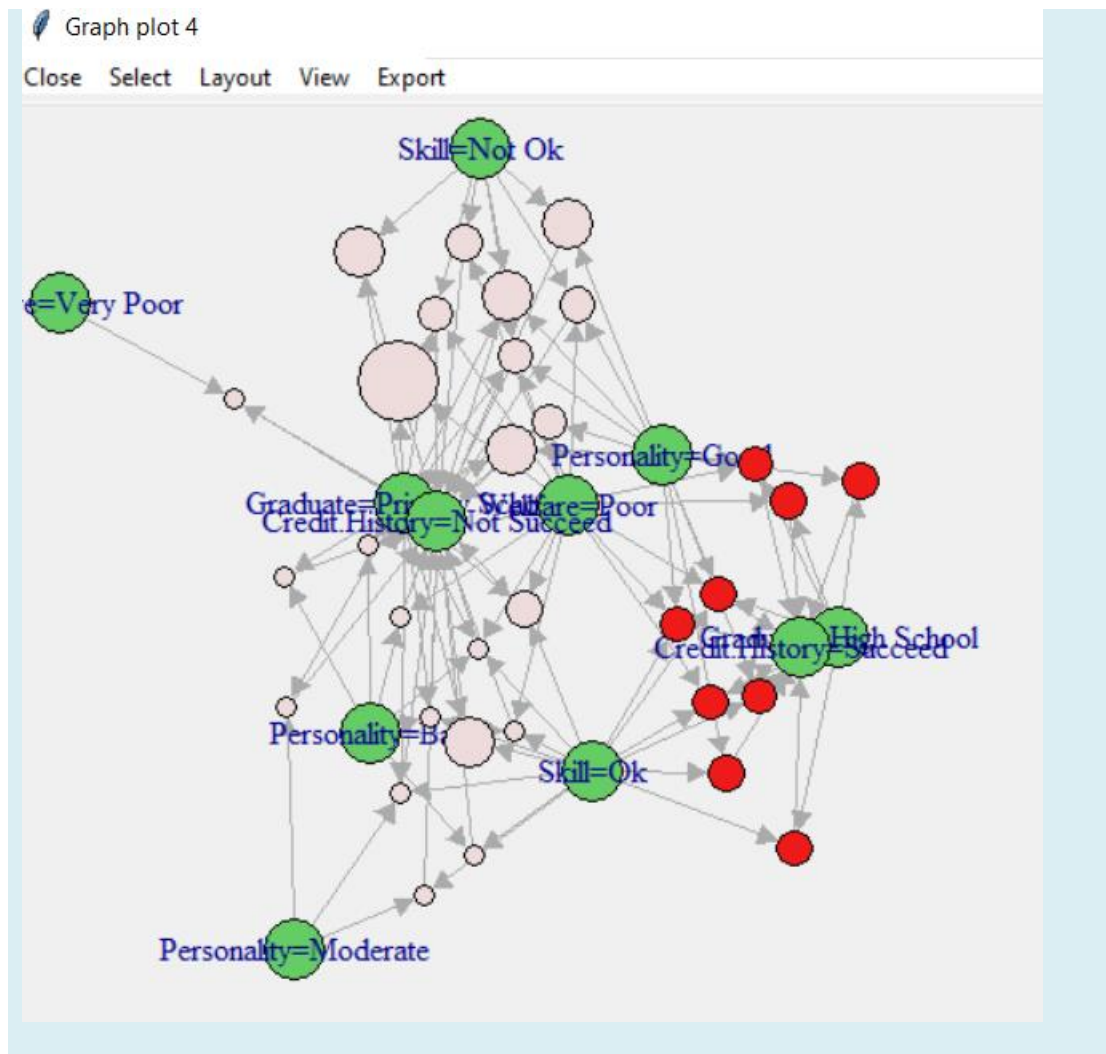
```

> dataibnu
  No Graduate Skill Wellfare Personality Credit.History
1  1  High School   Ok    Poor        Good      Succeed
2  2 Primary School Not Ok  Very Poor    Good    Not Succeed
3  3 Primary School   Ok    Poor    Moderate Not Succeed
4  4  High School   Ok    Poor        Good      Succeed
5  5 Primary School Not Ok  Poor     Good    Not Succeed
6  6 Primary School   Ok  Very Poor    Moderate Not Succeed
7  7 Primary School   Ok    Poor        Bad    Not Succeed
8  8 Primary School Not Ok  Poor     Good    Not Succeed
9  9 Primary School Not Ok  Poor     Good    Not Succeed
10 10 Primary School   Ok    Poor        Bad    Not Succeed
11 11  High School   Ok    Poor     Good      Succeed
Warning message:
In rm(list = cmd, envir = .tkplot.env) : object 'tkp.3' not found
> categoric<- c("Graduate", "Skill", "Wellfare", "Personality", "Credit.History")
> transactions <- as(dataibnu[, categoric], "transactions")
Warning message:
Column(s) 1, 2, 3, 4, 5 not logical or factor. Applying default discretization (see '? discretizeDF').
> itemFrequencyPlot(transactions, topN = 5)
> |

```







CEK LIST

Elemen Kompetensi	No Latihan	Penyelesaian	
		Selesai	Tidak selesai
1	1.1.1	✓	
	1.1.2	✓	

FORM UMPAN BALIK

Elemen Kompetensi	Tingkat Kesulitan	Tingkat Ketertarikan	Waktu Penyelesaian dalam menit
Memahami cara implementasi Association Rule di R	<input type="checkbox"/> Sangat Mudah <input type="checkbox"/> Mudah <input checked="" type="checkbox"/> Biasa <input type="checkbox"/> Sulit <input type="checkbox"/> Sangat Sulit	<input type="checkbox"/> Tidak Tertarik <input type="checkbox"/> Cukup Tertarik <input type="checkbox"/> Tertarik <input checked="" type="checkbox"/> Sangat Tertarik	30