

BAB V

ANALISIS DATA MENGGUNAKAN WEKA

5.1 PRA PROSES DATA

Pada Tahap ini dilakukan analisa data dengan tools Weka. Persiapan proses data dimulai dari menyalin data mentah kedalam *Microsoft Excel*, penyeleksian berdasarkan merek item, transformasi data ke 35 kategori item, pelabelan data dan konversi pelabelan data menjadi format csv. Adapun proses persiapan data sebagai berikut:

5.1.1 Data Eksperimen

Data mentah ini merupakan data transaksi penjualan pada Toko Novi sebelum dilakukannya penyeleksian data. Adapun data tersebut dapat dilihat pada gambar 5.1

A	B
1	minyak bimoli, daia, pepsodent, sabun nuvo, sukro,tictac, izzi, permen kopiko, protex
2	minyak tanah, jerigen kecil
3	lampu philip 5 wat
4	kopi kapal api, araya
5	pilus garuda, indomie rendang, telur ayam, gula 1 kg, tepung segitiga biru
6	aqua, sampoerna
7	kacang hijau, kara sun, gula 1 kg, tepung segitiga biru, palmia
8	sedotan warna
9	jarum pentul
10	terpal
11	masako, minyak rose brand, intermie, sarimie isi 2, supermi, telur
12	pisau silet, gillette
13	leter l , leter t, lem aicha-aibon
14	paku triplek, paku atap
15	ekonomi, sunlight, super pell,pasta gigi ciptadent,telur, blue band, bukrim, bahun jag
16	hatari orange, padimas
17	telur, minyak sunco, gula 1 kg, rokok surya, supermi ayam bawang, kopi bubuk AAA
18	benang jahit, jarum jahit

Gambar 5.1 Data Mentah

5.1.2 Seleksi Data

5.1.2.1 Seleksi Data Berdasarkan Merek *Item*

Pada tahap ini dilakukan penyeleksian data berdasarkan merek – merek dari *item* yang diambil. Data tersebut dapat dilihat pada gambar 5.2.

1	minyak bimoli, daia, pepsodent, sabun nuvo, sukro,tictac, izzi, permen kopiko, protex, bihun jagung
2	kopi kapal api, araya
3	pilus garuda, indomie rendang, telur ayam, gula 1 kg, tepung segitiga biru
4	aqua, sampoerna
5	kacang hijau, kara sun, gula 1 kg, tepung segitiga biru, palmia
6	sedotan warna
7	masako, minyak rose brand, intermie, sarimie isi 2, supermi, telur
8	ekonomi, sunlight, super pell, pasta gigi ciptadent, telur, blue band, bukrim, bihun jagung, minyak sunco
9	hatari orange, padimas
10	telur, minyak sunco, gula 1 kg, rokok surya, supermi ayam bawang, kopi bubuk AAA
11	kuaci, pop mie, aqua
12	molto ultra saset, sikat gigi formula, sabun lux, gula 1 kg, mama lemon, salisil, mie sedap bakso, intermie
13	kukubuma ener-G, surya
14	mie sedap goreng, supermie, teh bendera, minyak sovia, parfum izzi
15	vanile, pasta pandan, tepung maizena, telur, palmia, mie telur,

Gambar 5.2 Penyeleksian Data Mentah Berdasarkan Merek *Item*

5.1.2.2 Transformasi Data Ke 35 Kategori *Item*

Setelah dilakukan penyeleksian berdasarkan merek *item*, kemudian data tersebut ditransformasikan ke 35 kategori *item* yaitu mie, roti, minuman, tepung, margarin, parfum, shampo, pembalut, tisu, permen, rokok, plastik, bumbu, kecap, saos, kacang, teh, kopi, kerupuk, gula, kosmetik, pewangi, sabun, detergen, pasta gigi, minyak sayur, bahan kue, susu, telur, makanan ringan, pampers, obat, obat nyamuk, sikat gigi dan alat pecah belah. Adapun data tersebut dapat dilihat pada gambar 5.3:

1	minyak sayur, detergen, pasta gigi, sabun, makanan ringan, parfum, permen, pembalut, mie
2	kopi, rokok
3	makanan ringan, mie, telur, gula, tepung,
4	minuman, rokok
5	kacang, bumbu, gula, tepung, margarin
6	plastik
7	bumbu, minyak sayur, mie, telur
8	sabun, pasta gigi, telur, margarin, detergen, mie, minyak sayur
9	roti, makanan ringan
10	telur, minyak sayur, gula, rokok, mie, kopi
11	makanan ringan, mie, minuman
12	pewangi, sikat gigi, sabun, gula, obat mie
13	minuman, rokok
14	mie, teh, minyak sayur, parfum
15	bahan kue, tepung, telur, margarin, mie,
16	pempers, roti, mie, telur

Gambar 5.3 Transformasi Data Ke 35 Kategori Item

5.2 PELABELAN DATA

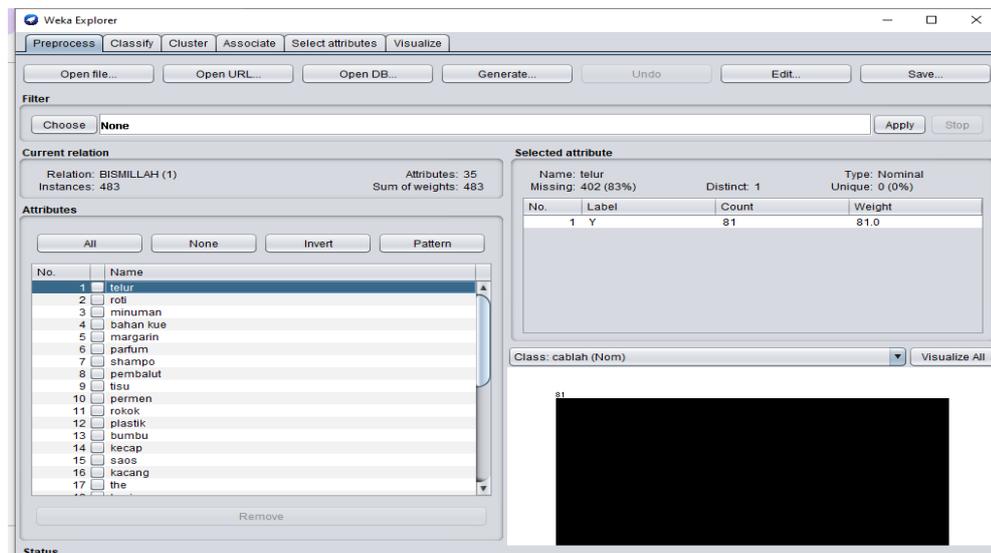
Pada tahapan ini dilakukan pelabelan data pada data yang telah diseleksi.

Adapun data tersebut dapat dilihat pada gambar 5.4 :

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	telur	roti	minuman	bahan kue	margarin	parfum	shampo	pembalut	tisu	permen	rokok	plastik	bumbu	kecap	saos	kacang	the
2	?	?	?	?	?	Y	?	Y	?	Y	?	?	?	?	?	?	?
3	?	?	?	?	?	?	?	?	?	?	Y	?	?	?	?	?	?
4	Y	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
5	?	?	Y	?	?	?	?	?	?	?	Y	?	?	?	?	?	?
6	?	?	?	?	Y	?	?	?	?	?	?	?	Y	?	?	Y	?
7	?	?	?	?	?	?	?	?	?	?	?	Y	?	?	?	?	?
8	Y	?	?	?	?	?	?	?	?	?	?	?	Y	?	?	?	?
9	Y	?	?	?	Y	?	?	?	?	?	?	?	?	?	?	?	?
10	?	Y	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
11	Y	?	?	?	?	?	?	?	?	?	Y	?	?	?	?	?	?
12	?	?	Y	?	?	?	?	?	?	?	?	?	?	?	?	?	?
13	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

Gambar 5.4 Pelabelan Data WEKA

Dari gambar 5.3, terdapat 35 atribut yang merupakan nama item yang telah dikategorikan. Terdapat 2 label yaitu true (Y) dan false (?). True (Y) jika nama item

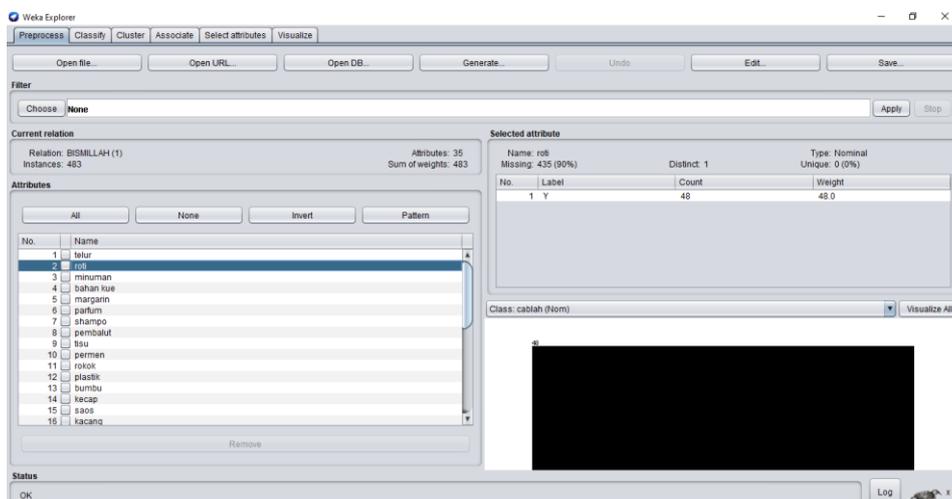


Gambar 5.6 Visualisasi Atribut Telur

Gambar 5.6 adalah visualisasi dari Atribut Telur. Diketahui bahwa dari 483 transaksi terdapat *missing* data sebanyak 402 atau 83%. Maka dapat dijelaskan bahwa jumlah atribut mie adalah 81.

2. Visualisasi Atribut Roti

Bentuk visualisasi atribut Roti dengan menggunakan tools *WEKA*, dapat dilihat pada gambar 5.7 :

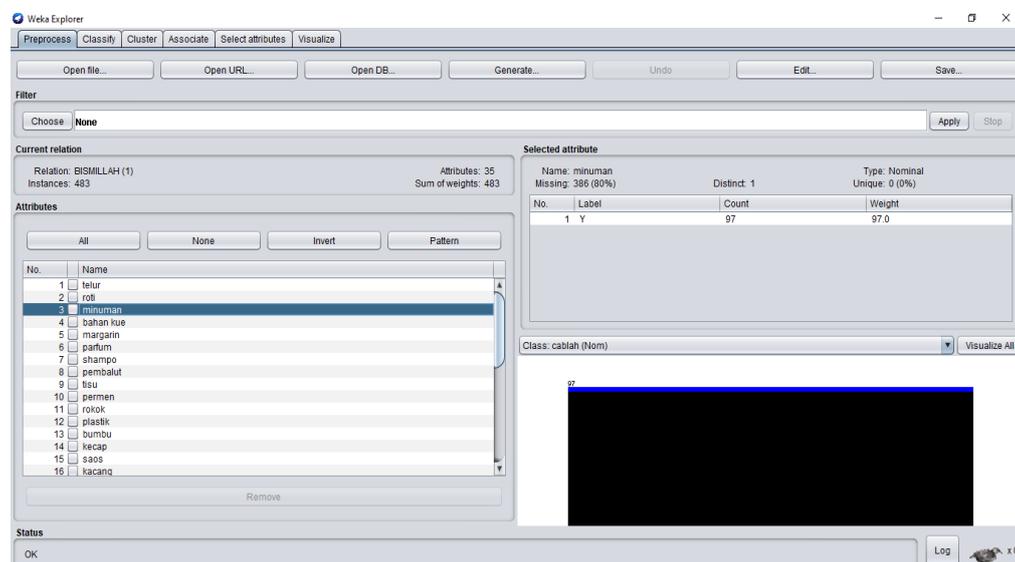


Gambar 5.7 Visualisasi Atribut Roti

Gambar 5.7 adalah visualisasi dari atribut Roti. Diketahui bahwa dari 483 transaksi terdapat *missing* data sebanyak 435 atau 90%. Maka dapat dijelaskan bahwa jumlah atribut roti adalah 48.

3. Visualisasi Atribut Minuman

Bentuk visualisasi atribut Minuman dengan menggunakan tools *WEKA*, dapat dilihat pada gambar 5.8 :

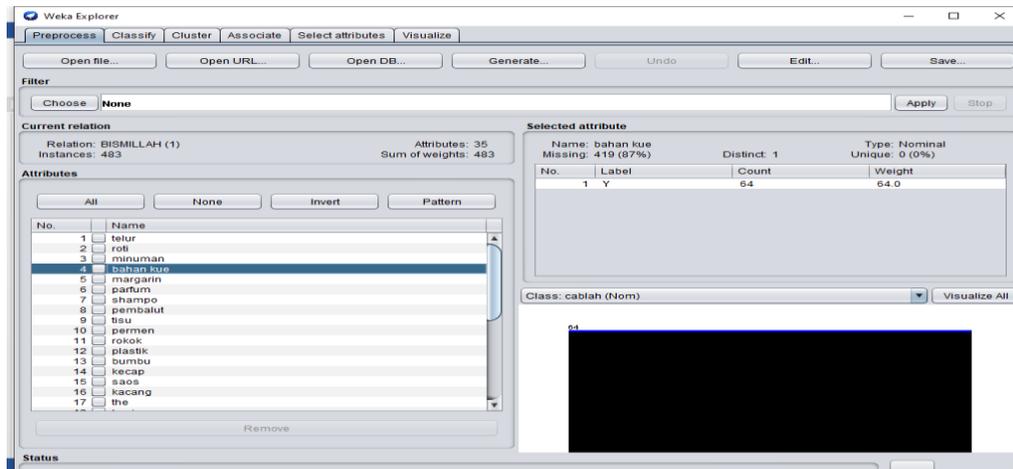


Gambar 5.8 Visualisasi Atribut Minuman

Gambar 5.8 adalah visualisasi dari Atribut Minuman. Diketahui bahwa dari 483 transaksi terdapat *missing* data sebanyak 386 atau 80%. Maka dapat dijelaskan bahwa jumlah atribut minuman adalah 97.

4. Visualisasi Atribut Bahan Kue

Bentuk visualisasi atribut Bahan Kue dengan menggunakan tools *WEKA*, dapat dilihat pada gambar 5.9 :

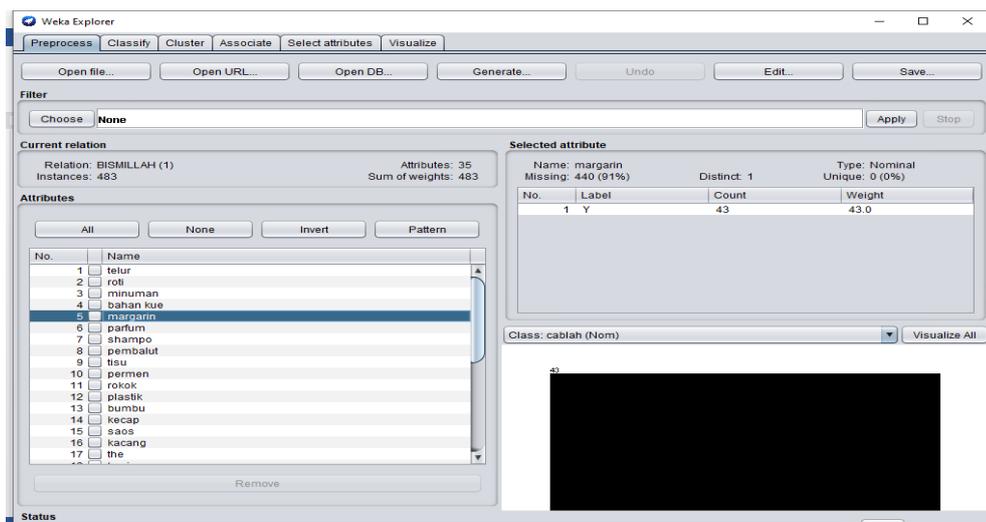


Gambar 5.9 Visualisasi Atribut Bahan Kue

Gambar 5.9 adalah visualisasi dari Atribut Bahan Kue. Diketahui bahwa dari 483 transaksi terdapat *missing* data sebanyak 419 atau 87%. Maka dapat dijelaskan bahwa jumlah atribut Bahan Kue adalah 64.

5. Visualisasi Atribut Margarin

Bentuk visualisasi Atribut Margarin dengan menggunakan tools WEKA, dapat dilihat pada gambar 5. 10:



Gambar 5.10 Visualisasi Atribut Margarin

Gambar 5.10 adalah visualisasi dari Atribut Margarin. Diketahui bahwa dari 483 transaksi terdapat *missing* data sebanyak 440 atau 91 %. Maka dapat dijelaskan bahwa jumlah atribut margarin adalah 43.

Pada tahapan selanjutnya visualisasi atribut akan ditampilkan dalam bentuk tabulasi. Tabulasi yang ditampilkan adalah *missing data count* (jumlah) data, dapat dilihat pada tabel 5.1 :

Tabel 5.1 Hasil Visualisasi Setiap Atribut Dengan Menggunakan WEKA

No	Nama Atribut	<i>Missing Data</i>	Jumlah Atribut
1.	Parfum	471 (98%)	12
2.	Shampo	442 (92%)	41
3.	Pembalut	456 (94%)	27
4.	Tisu	478 (99%)	5
5.	Permen	423 (88%)	60
6.	Rokok	406 (84%)	77
7.	Plastik	469 (97%)	14
8.	Bumbu	380 (79%)	103
9.	Kecap	472 (98%)	11
10.	Saos	462 (96%)	21
11.	Kacang	466 (96%)	17
12.	Teh	445 (92%)	38
13.	Kopi	436 (90%)	47
14.	Kerupuk	455 (94%)	28

15.	Gula	414 (86%)	69
16.	Kosmetik	458 (95%)	25
17.	Pewangi	448 (93%)	35
18.	Sabun	404 (84%)	79
19.	Detergen	408 (84%)	75
20.	Pasta gigi	452 (94%)	31
21.	Minyak sayur	371 (77%)	112
22.	Tepung	411 (85%)	72
23.	Susu	432 (89%)	51
24.	Mie	364 (75%)	119
25.	Makanan ringan	374 (77%)	109
26.	Pampers	470 (97%)	13
27.	Obat	439 (91%)	44
28.	Obat nyamuk	469 (97%)	14
29.	Sikat gigi	475 (98%)	8
30.	Alat pecah belah	464 (96%)	19

5.4 HASIL ANALISIS ASOSIASI DENGAN TOOLS WEKA

Pada tahapan ini menampilkan hasil analisis asosiasi data dengan tools *WEKA*. Hasil analisis asosiasi dapat dilihat pada gambar 5.11 :

```

ssociator output
mie
makanan ringan
pempers
obat
obat nyamuk
sikat gigi
cablah
=== Associator model (full training set) ===

Apriori
=====

Minimum support: 0.05 (24 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 19

Generated sets of large itemsets:

Size of set of large itemsets L(1): 25
Size of set of large itemsets L(2): 18

Best rules found:

1. telur=Y 81 ==> mie=Y 48 <conf:(0.59)> lift:(2.41) lev:(0.06) [28] conv:(1.8)
2. margarin=Y 43 ==> tepung=Y 25 <conf:(0.58)> lift:(3.9) lev:(0.04) [18] conv:(1.93)
3. bahan kue=Y 64 ==> tepung=Y 34 <conf:(0.53)> lift:(3.56) lev:(0.05) [24] conv:(1.76)

```

Gambar 5.11 Analisis Asosiasi

Pada gambar 5.11 merupakan hasil analisis asosiasi data transaksi penjualan pada Toko Novi, dimana nilai minimum *supportnya* 0,05 dan nilai minimum *confidencenya* 0,5. Dari analisis tersebut menghasilkan rule terbaik, yaitu :

1. Jika membeli Telur maka membeli Mie dengan confidence 59%.
2. Jika membeli Margarin maka membeli Tepung dengan confidence 58%.
3. Jika membeli Bahan Kue maka membeli Tepung dengan confidence 53%.

5.5 PERBANDINGAN HASIL

Tahapan ini menampilkan hasil perbandingan data dengan menggunakan manual Excel dan Weka. Hasil perbandingan data tersebut dapat dilihat pada tabel 5.2 :

Tabel 5.2 Perbandingan Hasil

Rule yang dihasilkan menggunakan manual Excel		Manual Excel	Rule yang Dihasilkan menggunakan tools Weka		Rule yang dihasilkan menggunakan tools Weka
	<i>support</i>	<i>Confidence</i>		<i>support</i>	<i>confidence</i>
Telur, Mie	9%	59 %	Telur, Mie	9%	59 %
Margarin, Tepung	5%	58 %	Margarin, Tepung	5%	58 %
Bahan Kue, Tepung	7%	53%	Bahan Kue, Tepung	7%	53%

Tabel 5.2 menjelaskan tentang perbandingan hasil antara perhitungan manual, perhitungan menggunakan tools *WEKA*. Rule yang dihasilkan sama yaitu, jika membeli Telur maka membeli Mie dengan support 9% dan *Confidence* 59%, jika membeli Margarin maka membeli Tepung dengan support 5% dan *Confidence* 58%, jika membeli Bahan Kue maka membeli Tepung dengan support 7% dan *Confidence* 53%.