

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil

Berdasarkan penelitian yang telah dilakukan oleh penulis, penulis menemukan beberapa asosiasi ataupun hubungan yang terjadi antar data yang pemanfaatan data mining dengan menggunakan algoritma apriori tentu sangat membantu pembaca untuk menemukan pola hubungan antar dataset. Adapun asosiasi antar data tersebut adalah

4.1.1 Implementasi Rapid Miner

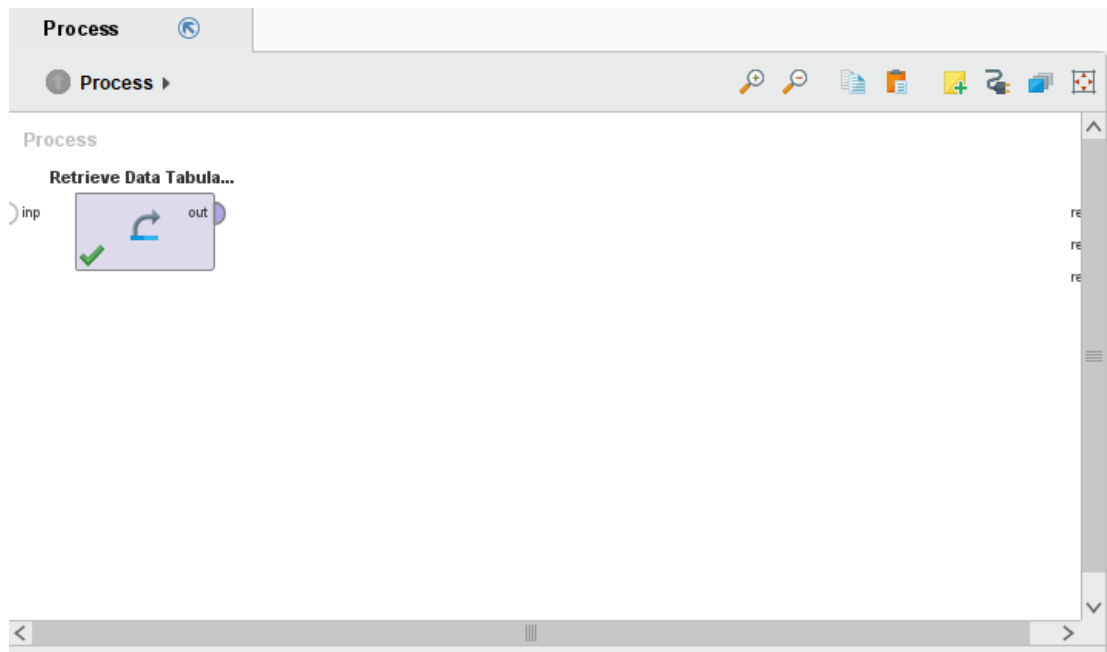
Pada bagian ini adalah mengimplementasikan Algoritma Apriori menggunakan perangkat lunak Rapid Miner (9.8.0.001). Desain proses yang diimplementasikan dalam rapid miner untuk menentukan hubungan relasi menggunakan *retrive, replace missing value, discretize, FP-Growth dan create association*.

Dibawah ini merupakan gambaran proses asosiasi data tracer study pada IIB Darmajaya.

Sebelum masuk kedalam import data, data yang kita simpan yaitu data excel yang sudah di cleaning secara manual dengan excel dan juga menggunakan *replace missing value* agar lebih maksimal dalam pengolahannya setelah data sudah dibersihkan langkah pertama kita dapat mengimport data yang telah disiapkan sebelumnya ke dalam rapid miner.

4.1.1.1 Import Data

Seperti dijelaskan digambar 4.1 ini adalah proses retrieve data ke dalam rapid miner yaitu proses memasukan atau meng-*import* data dari folder yang disimpan lalu dimasukan atau ditarik kedalam proses rapid miner.



Gambar 4. 1 Proses Retrive Data ke Dalam Rapid Miner

Retrieve data tabular adalah nama folder yang sudah di import dari folder yang isinya berupa data-data berupa data excel yang sudah diolah kemudian ditarik atau dimasukkan kedalam *view process* atau lembar proses.

4.1.1.2 Tampilan Data Import

Ini adalah tampilan data yang telah di *import* data-data inilah hasil dari data yang sudah dijalankan dari *import* data seperti digambar 4.1.

Row No.	Laki Laki	Perempuan	<2-3	>3 -3,5	>3,5	Tepat	Tidak Tepat	1-6 Bulan	7-18 Bulan	>18 Bulan
1	1	0	0	0	1	1	0	0	0	0
2	0	1	0	1	0	1	0	1	0	0
3	1	0	0	1	0	1	0	1	0	0
4	1	0	1	0	0	0	1	1	0	0
5	1	0	0	1	0	0	1	0	1	0
6	1	0	0	1	0	1	0	0	0	0
7	1	0	0	1	0	1	0	1	0	0
8	1	0	0	1	0	1	0	1	0	0
9	1	0	0	1	0	1	0	1	0	0
10	1	0	0	0	1	1	0	0	0	0
11	0	1	0	1	0	1	0	0	1	0
12	1	0	0	1	0	1	0	1	0	0
13	1	0	0	1	0	1	0	1	0	0
14	1	0	0	1	0	1	0	0	0	0
15	1	0	0	1	0	1	0	0	0	0

Gambar 4. 2 Tampilan Data yang Telah di Import

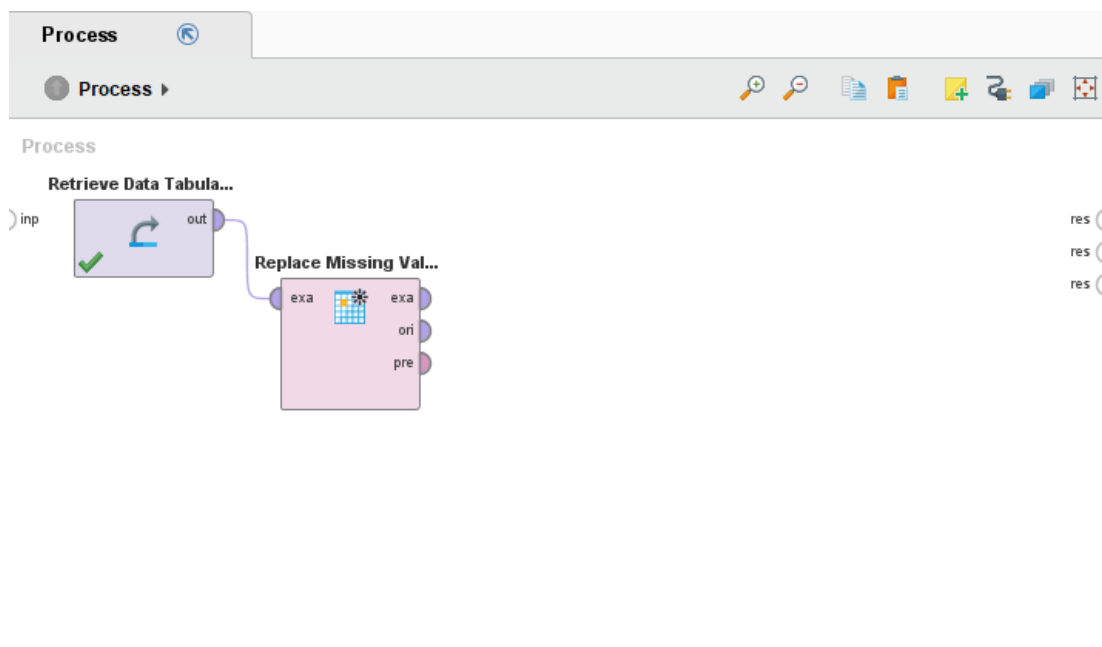
Terdapat beberapa data seperti data perempuan, pria, ipk <2-3, ipk 3-3,5 kemudian ipk >3,5, lulus tepat waktu dan tidak tepat waktu, masa tunggu kerja sebelum 6 bulan dan sesudah 6 bulan, gaji, pekerjaan dan bidang kerja.

Data yang telah berhasil di import adalah data yang kelak akan diolah oleh penulis untuk menemukan kandidat item set dan menemukan aturan asosiasi untuk menghasilkan pola tracer study yang nantinya akan memudahkan pihak IIB Darmajaya untuk menemukan, mengawasi serta mengevaluasi alumni. Informasi ini kelak akan menjadi sumber informasi dan untuk mengetahui penyerapan, proses, dan posisi lulusan dalam dunia kerja serta membantu program pemerintah dalam memetakan kebutuhan dunia kerja dengan kompetensi yang diperoleh dari perguruan tinggi.

Selanjutnya, setelah mengimport data kita data membersihkan data yang mengalami kerusakan, kehilangan dan memeriksa data inkonsisten. Untuk melakukan cleansing data tersebut penulis menggunakan *replace missing value*.

4.1.1.3 Replace Missing Value

Pada gambar 4.3 terdapat *replace missing value* yaitu proses penyaringan data yang noise atau data yang double sehingga data itu bisa disaring sehingga tidak ada lagi data yang noise.



Gambar 4. 3 Replace Missing Value

Sesudah data excel dimasukkan seperti Gambar 4.1 selanjutnya ialah memasukkan *replace missing value* di proses rapid miner agar membersihkan data yang kosong atau sama.

4.1.1.4 Data yang sudah di Replace *Missing Value*

Tampilan gambar 4.4 ini merupakan tampilan dari *replace missing value* yang sudah selesai dilakukan. Selanjutnya penulis menggunakan *discritze by frequency* dan *numerical to binominal* kedalam field.

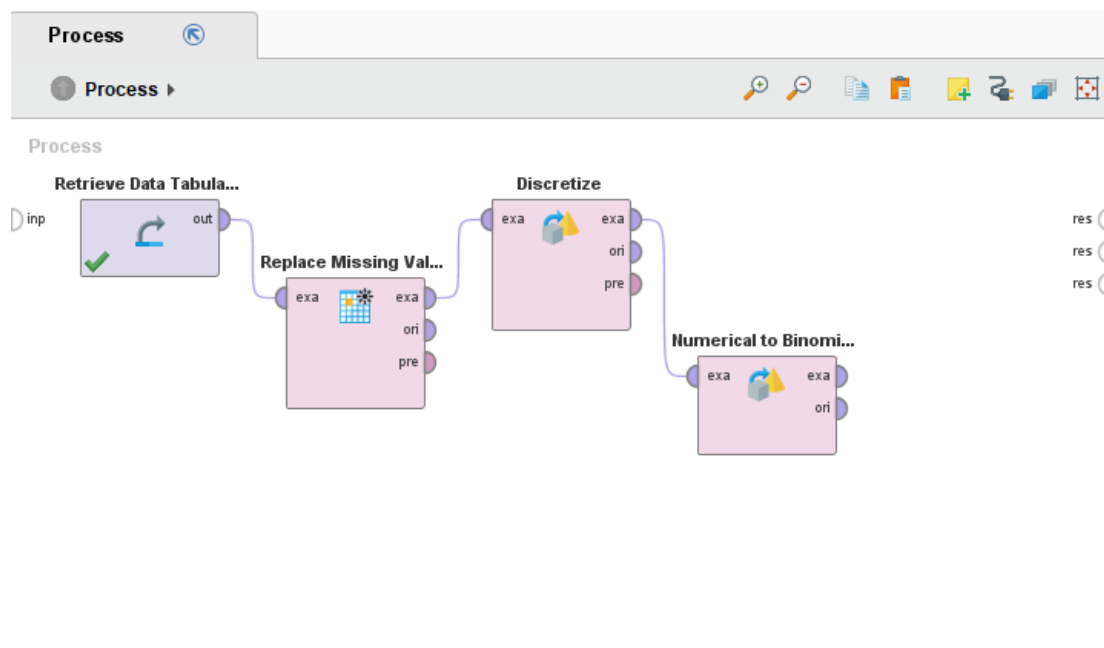
Row No.	Laki Laki	Perempuan	<2,3	>3 -3,5	>3,5	Tepat	Tidak Tepat	1-6 Bulan	7-18 Bulan	>18 Bulan
1	1	0	0	0	1	1	0	0	0	0
2	0	1	0	1	0	1	0	1	0	0
3	1	0	0	1	0	1	0	1	0	0
4	1	0	1	0	0	0	1	1	0	0
5	1	0	0	1	0	0	1	0	1	0
6	1	0	0	1	0	1	0	0	0	0
7	1	0	0	1	0	1	0	1	0	0
8	1	0	0	1	0	1	0	1	0	0
9	1	0	0	1	0	1	0	1	0	0
10	1	0	0	0	1	1	0	0	0	0
11	0	1	0	1	0	1	0	0	1	0
12	1	0	0	1	0	1	0	1	0	0
13	1	0	0	1	0	1	0	1	0	0
14	1	0	0	1	0	1	0	0	0	0
15	1	0	0	1	0	1	0	0	0	0

Gambar 4. 4 Data yang telah dilakukan Replace Missing Value

Seperti yang dilihat ini adalah tampilan yang sudah dilakukan oleh operator *replace missing value* data yang double atau kosong sudah tidak ada lagi karena operator ini menghilangkan data tersebut.

4.1.1.5 *Descritize* dan *Numerical to Binominal*

Selanjutnya kita dapat memasukan *Descritize* dan *Numerical to Binominal* untuk memasukan algoritma apriori dan menambahkan operator *Create Association Rules*.

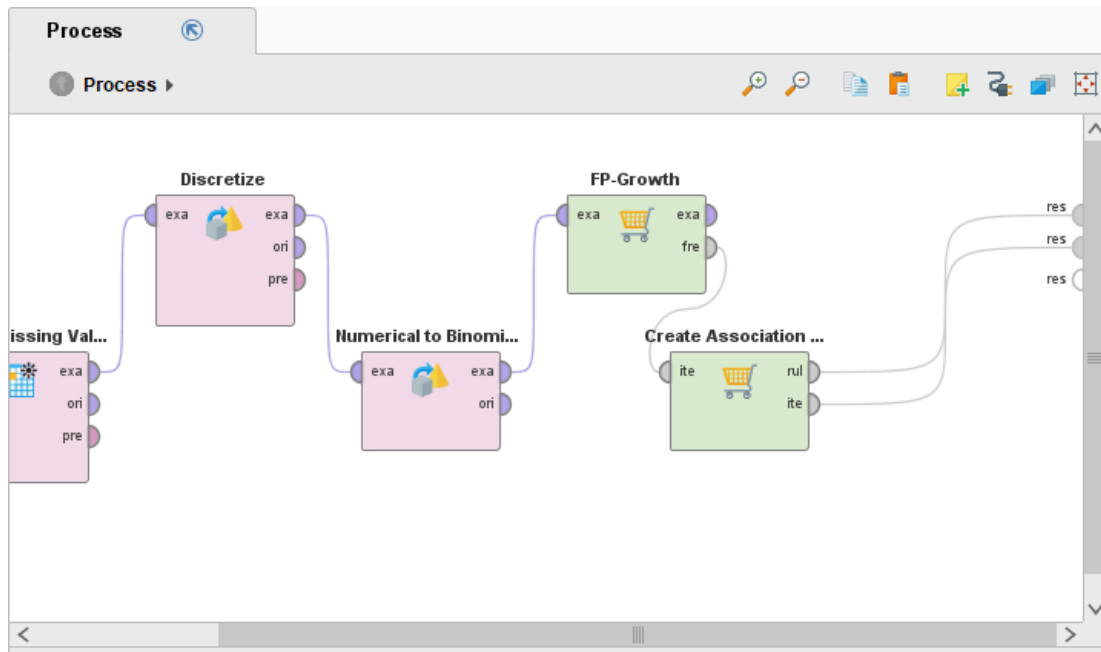


Gambar 4. 5 Proses Penambahan Operator *Discretize* dan *Numerical to Binominal*

Tahapan selanjutnya memasukan operator *Descritize* dan *FP-Growth* tentunya untuk menyesuaikan pada tahapan sebelumnya sesudah *replace missing value*.

4.1.1.6 Menjalankan Apriori

Memasukkan FP-Growth dan *Create Association Rules* pada lembar kerja atau *view process*.

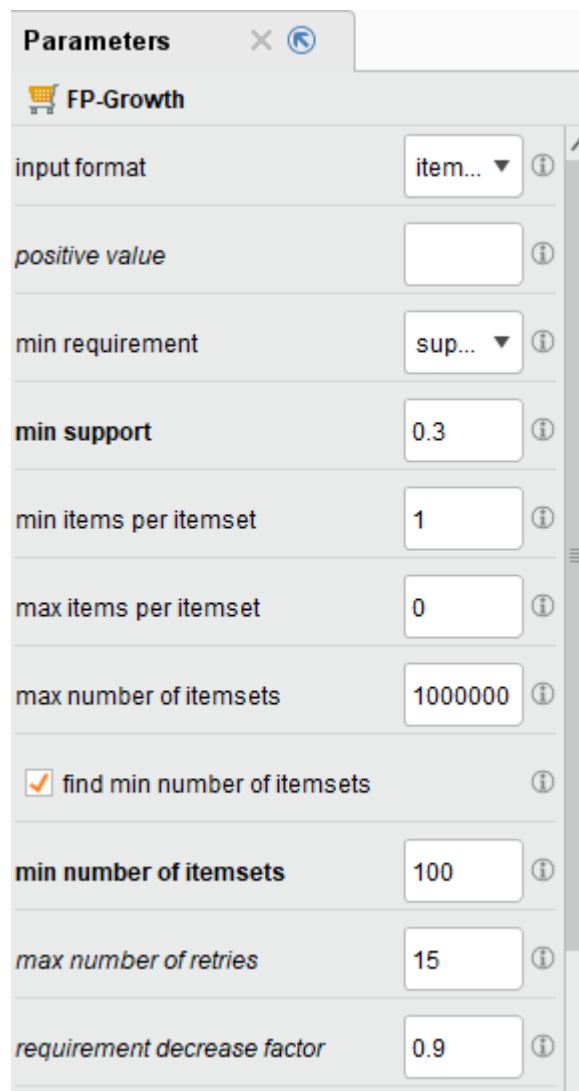


Gambar 4. 6 Proses Membubuhkan Algoritma Apriori

Setelah semua sudah dimasukkan lalu dijalankan dengan menekan tombol play pada menu play button yang terdapat diatas atau bisa langsung menekan kunci F11 pada keyboard.

4.1.1.7 Minimum Support

Untuk menentukan *minimum support* dan *minimum confidence* dilakukan secara manual pada operator *FP-growth* dan *Create Association Rules* dalam bentuk pecahan decimal. Dengan ketentuan *minimum support* 3% dan *maximum confidence* sebesar 30%.



The image shows a 'Parameters' dialog box for the 'FP-Growth' algorithm. The 'min support' parameter is highlighted in bold and set to the value 0.3. Other parameters include 'input format' (item...), 'positive value', 'min requirement' (sup...), 'min items per itemset' (1), 'max items per itemset' (0), 'max number of itemsets' (1000000), a checked 'find min number of itemsets' checkbox, 'min number of itemsets' (100), 'max number of retries' (15), and 'requirement decrease factor' (0.9).

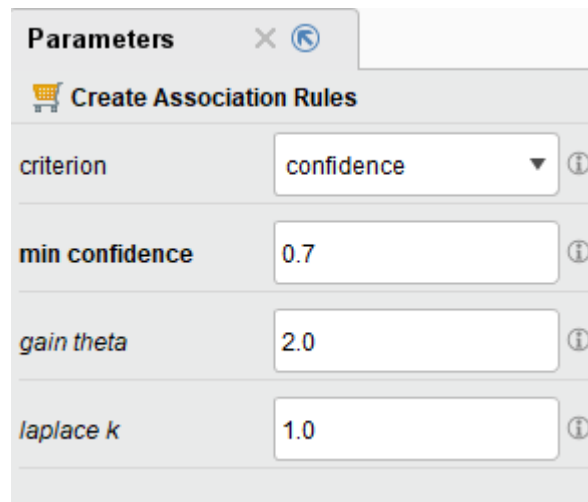
Parameter	Value
input format	item...
positive value	
min requirement	sup...
min support	0.3
min items per itemset	1
max items per itemset	0
max number of itemsets	1000000
<input checked="" type="checkbox"/> find min number of itemsets	
min number of itemsets	100
max number of retries	15
requirement decrease factor	0.9

Gambar 4. 7 Menentukan Minimum support

Penentuan minimum support dilakukan dengan menggunakan rumus yang telah dijabarkan pada Bab 2 yang berguna sebagai batasan frekuensi kejadian atau support count yang harus dipenuhi suatu kelompok data untuk dapat dijadikan aturan. Pada penelitian ini menggunakan minimal support sebesar 0,3 atau 30%.

4.1.1.8 Minimum Confidence

Memasukkan minimum confidence yang sudah dilakukan secara manual oleh FP-Growth dan Create Association Rules dalam bentuk pecahan desimal.



Parameter	Value
criterion	confidence
min confidence	0.7
gain theta	2.0
laplace k	1.0

Gambar 4. 8 Menentukan Minimum confidence

Minimum confidence yaitu parameter yang mendefinisikan minimum level dari confidence yang harus dipenuhi oleh aturan yang berkualitas. Minimum ini harus dipenuhi oleh aturan-aturan yang berkualitas dan harus terpenuhi oleh suatu kelompok untuk dijadikan sebuah aturan. Minimal confidence yang digunakan pada penelitian ini yaitu, 0.7 atau sebesar 70%.

4.1.1.9 Itemset

Dengan hasil sebagai berikut, pada gambar 4.9 ini adalah itemset yang dihasilkan dari proses rapid miner yang dilakukan.

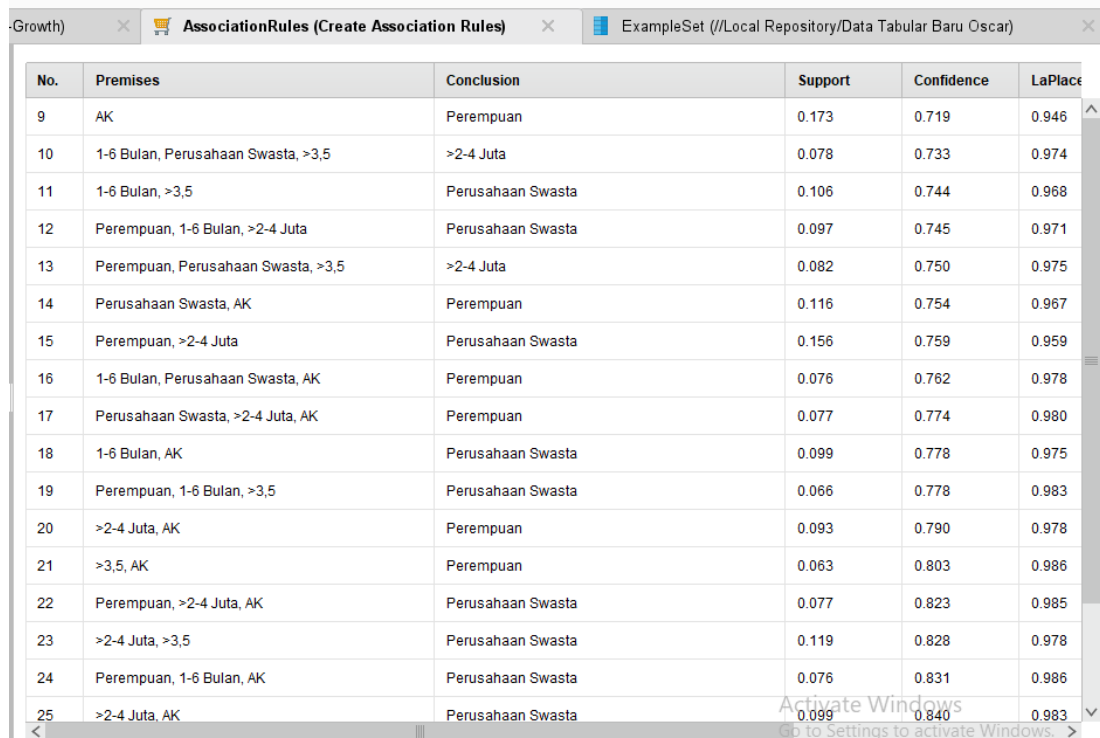
Size	Support	Item 1	Item 2	Item 3	Item 4
1	0.488	Perempuan			
1	0.470	1-6 Bulan			
1	0.467	Perusahaan Swasta			
1	0.409	Manajemen			
1	0.400	>2-4 Juta			
1	0.303	>3,5			
1	0.241	Tidak Bekerja			
1	0.240	AK			
1	0.193	TI			
1	0.190	0-2 Juta			
1	0.171	Tidak Tepat			
1	0.165	>4 Juta			
1	0.142	Wiraswasta			
1	0.134	Instansi Pemerintah			
1	0.130	SI			
1	0.122	<-2-3			
2	0.212	Perempuan	1-6 Bulan		

Gambar 4. 9 Itemset

Ada beberapa *itemset* pada gambar 4.9 ada 94 *itemset* dengan total *max size* nya adalah 4 seperti contoh perempuan memiliki *support* 0,488%, kemudian masa tunggu kerja 1-6 bulan memiliki *support* 0,470% dan lain-lain.

4.1.1.10 Asosiasi

Ini adalah tabel asosiasi yang didapatkan dari proses yang dijalankan di *rapid miner*



No.	Premises	Conclusion	Support	Confidence	LaPlace
9	AK	Perempuan	0.173	0.719	0.946
10	1-6 Bulan, Perusahaan Swasta, >3,5	>2-4 Juta	0.078	0.733	0.974
11	1-6 Bulan, >3,5	Perusahaan Swasta	0.106	0.744	0.968
12	Perempuan, 1-6 Bulan, >2-4 Juta	Perusahaan Swasta	0.097	0.745	0.971
13	Perempuan, Perusahaan Swasta, >3,5	>2-4 Juta	0.082	0.750	0.975
14	Perusahaan Swasta, AK	Perempuan	0.116	0.754	0.967
15	Perempuan, >2-4 Juta	Perusahaan Swasta	0.156	0.759	0.959
16	1-6 Bulan, Perusahaan Swasta, AK	Perempuan	0.076	0.762	0.978
17	Perusahaan Swasta, >2-4 Juta, AK	Perempuan	0.077	0.774	0.980
18	1-6 Bulan, AK	Perusahaan Swasta	0.099	0.778	0.975
19	Perempuan, 1-6 Bulan, >3,5	Perusahaan Swasta	0.066	0.778	0.983
20	>2-4 Juta, AK	Perempuan	0.093	0.790	0.978
21	>3,5, AK	Perempuan	0.063	0.803	0.986
22	Perempuan, >2-4 Juta, AK	Perusahaan Swasta	0.077	0.823	0.985
23	>2-4 Juta, >3,5	Perusahaan Swasta	0.119	0.828	0.978
24	Perempuan, 1-6 Bulan, AK	Perusahaan Swasta	0.076	0.831	0.986
25	>2-4 Juta, AK	Perusahaan Swasta	0.099	0.840	0.983

Gambar 4. 10 Tabel Asosiasi

Tabel asosiasi pada *create association rules* terlihat beberapa seperti jenis kelamin perempuan dengan IPK >3,5 maka memiliki kecenderungan 75% untuk bekerja di perusahaan swasta dan rule ini mewakili 82% dari seluruh data yang ada dan lain-lain seperti pada tampilan gambar diatas.

Sehingga dapat dilihat, terdapat beberapa aturan asosiasi yang dapat dibentuk. Keputusan ini kelak dapat membantu pihak IIB Darmajaya untuk melacak jejak alumni pada tahun 2019 dan 2020.

Hasil pembentukan aturan asosiasi dari pengujian pada aplikasi RapidMiner Studio dapat dilihat pada gambar berikut:

Association Rules

```

Association Rules
[>2-4 Juta, AK] --> [1-6 Bulan] (confidence: 0.700)
[1-6 Bulan, >2-4 Juta] --> [Perusahaan Swasta] (confidence: 0.705)
[Wiraswasta] --> [1-6 Bulan] (confidence: 0.708)
[>2-4 Juta] --> [Perusahaan Swasta] (confidence: 0.710)
[Perusahaan Swasta, >3,5] --> [>2-4 Juta] (confidence: 0.711)
[1-6 Bulan, AK] --> [Perempuan] (confidence: 0.713)
[1-6 Bulan, Perusahaan Swasta, AK] --> [>2-4 Juta] (confidence: 0.714)
[Perusahaan Swasta, >2-4 Juta, AK] --> [1-6 Bulan] (confidence: 0.714)
[AK] --> [Perempuan] (confidence: 0.719)
[1-6 Bulan, Perusahaan Swasta, >3,5] --> [>2-4 Juta] (confidence: 0.733)
[1-6 Bulan, >3,5] --> [Perusahaan Swasta] (confidence: 0.744)
[Perempuan, 1-6 Bulan, >2-4 Juta] --> [Perusahaan Swasta] (confidence: 0.745)
[Perempuan, Perusahaan Swasta, >3,5] --> [>2-4 Juta] (confidence: 0.750)
[Perusahaan Swasta, AK] --> [Perempuan] (confidence: 0.754)
[Perempuan, >2-4 Juta] --> [Perusahaan Swasta] (confidence: 0.759)
[1-6 Bulan, Perusahaan Swasta, AK] --> [Perempuan] (confidence: 0.762)
[Perusahaan Swasta, >2-4 Juta, AK] --> [Perempuan] (confidence: 0.774)
[1-6 Bulan, AK] --> [Perusahaan Swasta] (confidence: 0.778)
[Perempuan, 1-6 Bulan, >3,5] --> [Perusahaan Swasta] (confidence: 0.778)
[>2-4 Juta, AK] --> [Perempuan] (confidence: 0.790)
[>3,5, AK] --> [Perempuan] (confidence: 0.803)
[Perempuan, >2-4 Juta, AK] --> [Perusahaan Swasta] (confidence: 0.823)
[>2-4 Juta, >3,5] --> [Perusahaan Swasta] (confidence: 0.828)
[Perempuan, 1-6 Bulan, AK] --> [Perusahaan Swasta] (confidence: 0.831)
[>2-4 Juta, AK] --> [Perusahaan Swasta] (confidence: 0.840)
[1-6 Bulan, >2-4 Juta, >3,5] --> [Perusahaan Swasta] (confidence: 0.857)
[1-6 Bulan, >2-4 Juta, AK] --> [Perusahaan Swasta] (confidence: 0.857)
[Perempuan, >2-4 Juta, >3,5] --> [Perusahaan Swasta] (confidence: 0.863)

```

Gambar 3. 2 Association Rules

4.2 Analisa Hasil

Penerapan metode asosiasi dan algoritma apriori dalam mencari pola asosiatif dengan memanfaatkan data *tracer study* di Institut Informatika dan Bisnis Darmajaya. Dataset *tracer study* yang digunakan adalah tahun 2019-2020 dengan total 845 record data yang akan diuji pada skenario model pembentukan aturan asosiasi algoritma apriori. Melalui pengujian menggunakan aplikasi RapidMiner Studio penulis

menerapkan nilai *minimum support* sebesar 0,3 atau 30% dan *minimum confidence* juga sebesar 0,7 atau 70%.

Dalam tahapan pertama, proses perhitungan nilai *support* terhadap masing-masing item dan dari 25 item yang memenuhi syarat terhadap *minimum support* yang ditentukan, yakni Laki-Laki dengan nilai *support* 51,24%, selanjutnya Perempuan dengan nilai *support* 48,75%, IPK >3-3,5 dengan nilai *support* 57,75%, kemudian Tepat waktu dengan *support* 82,95%, Masa Tunggu Kerja 1-6 Bulan nilai *support* 46,98%, Gaji 2-4 Juta nilai *support* 39,88%, item Staff nilai *support* 63,90%, Manajemen dengan nilai *support* 40,82% dan Perusahaan Swasta dengan nilai *support* 46,62%.

Support	Item 1
0.488	Perempuan
0.470	1-6 Bulan
0.467	Perusahaan Swasta
0.409	Manajemen
0.400	>2-4 Juta
0.303	>3,5

Gambar 4. 11 K-1 itemset nilai minimum support pada RapidMiner

Pada perhitungan *support* terhadap kombinasi k-2 itemset, pada pengujian ini yang memenuhi syarat dalam mencapai nilai *minimum support* yang ditentukan. Berikut adalah hasil dari pengujian pada aplikasi RapidMiner Studio.

0.212	Perempuan	1-6 Bulan
0.242	Perempuan	Perusahaan Swasta
0.207	Perempuan	Manajemen
0.206	Perempuan	>2-4 Juta
0.197	Perempuan	>3,5
0.124	Perempuan	Tidak Bekerja
0.173	Perempuan	AK
0.098	Perempuan	0-2 Juta
0.063	Perempuan	>4 Juta
0.283	1-6 Bulan	Perusahaan Swasta
0.162	1-6 Bulan	Manajemen
0.260	1-6 Bulan	>2-4 Juta
0.143	1-6 Bulan	>3,5

Gambar 4. 12 K-2 itemset nilai minimum support pada RapidMiner

Selanjutnya untuk melihat seberapa besarnya keterkaitan antara masing-masing item, di tahapan pengujian ini penulis menentukan nilai minimum confidence sebesar 0,7 atau 70% pada aplikasi RapidMiner Studio seperti gambar dibawah ini.

No.	Premises	Conclusion	Support	Confidence
10	1-6 Bulan, Perusahaan Swasta, >3,5	>2-4 Juta	0.078	0.733
11	1-6 Bulan, >3,5	Perusahaan Swasta	0.106	0.744
12	Perempuan, 1-6 Bulan, >2-4 Juta	Perusahaan Swasta	0.097	0.745
13	Perempuan, Perusahaan Swasta, >3,5	>2-4 Juta	0.082	0.750
14	Perusahaan Swasta, AK	Perempuan	0.116	0.754
15	Perempuan, >2-4 Juta	Perusahaan Swasta	0.156	0.759
16	1-6 Bulan, Perusahaan Swasta, AK	Perempuan	0.076	0.762
17	Perusahaan Swasta, >2-4 Juta, AK	Perempuan	0.077	0.774
18	1-6 Bulan, AK	Perusahaan Swasta	0.099	0.778
19	Perempuan, 1-6 Bulan, >3,5	Perusahaan Swasta	0.066	0.778
20	>2-4 Juta, AK	Perempuan	0.093	0.790
21	>3,5, AK	Perempuan	0.063	0.803
22	Perempuan, >2-4 Juta, AK	Perusahaan Swasta	0.077	0.823
23	>2-4 Juta, >3,5	Perusahaan Swasta	0.119	0.828
24	Perempuan, 1-6 Bulan, AK	Perusahaan Swasta	0.076	0.831
25	>2-4 Juta, AK	Perusahaan Swasta	0.099	0.840
26	1-6 Bulan >2-4. Juta >3.5	Perusahaan Swasta	0.078	0.857

Gambar 4. 13 Aturan asosiasi

Ini adalah hasil dari proses Association Rules yang dijalankan RapidMiner Studio berdasarkan ketentuan yang sudah dilakukan dengan *minimum support* 0,3 atau 30% dan *minimum confidence* dengan nilai 0,7 atau 70%. Data yang diproses adalah data ditahun 2019-2020 yang berjumlah 845 data record dari *tracer study* Institut Informatika dan Bisnis Darmajaya .