

BAB III

METODE PENELITIAN

3.1. Data Transaksi

Dataset dalam suatu penelitian merupakan kumpulan individu atau obyek yang merupakan sifat sifat umum. Menurut Arikunto (2017) mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk . Maka dari penjelasan para ahli tersebut menetapkan data transaksi dalam penelitian ini adalah "data-data transaksi penjualan pulsa dan kuota di Gemilang Cell". Dan jumlah data yang terdapat dalam transaksi penjualan pulsa dan kuota di Gemilang Cell yakni mencapai 364 transaksi dari bulan 1 Mei 2019- 3 Agustus 2019 dan ini peneliti mengambil beberapa sampel data transaksi penjualan pulsa Untuk menentukan algoritma apriorinya dan aturan asosiasinya.

Tabel 3.1 Data Transaksi Penjualan

NO	Id Transaksi	Tgl Transaksi	Keterangan	Debet	Kredit	Saldo
1	11038695	08/05/2019 21:43:25	Simpati 50 ribu	49900.	0.00	375300.00
2	11038703	08/05/2019 21:46:55	Simpati 25 ribu	25350.0	0.00	349950.00
3	11040037	09/05/2019 12:27:42	Indosat 5 Ribu	6200.0	0.00	332550.00
4	11040039	09/05/2019 12:28:35	Indosat Sms 5 ribu	6500.0	0.00	326050.00
5	11043531	10/05/2019 16:41:11	Simpati 20 ribu	206500	0.00	940300.00
6	11043533	10/05/2019 16:41:37	Simpati 20 ribu	10800.0	0.00	929500.00
.						
.						
.						
.						
364	112666393	27/07/2019 00:16:08	Kuota Simpati 2GB	33650.00	0.00	1055900.00

3.2. Data Cleaning

Dalam nota transaksi penjualan terdapat beberapa atribut antara lain id transaksi, tanggal transaksi, kode produk, debet, kredit, saldo. Setelah dilakukan proses pembersihan atribut-atribut yang tersisa adalah id transaksi, tanggal transaksi, kode produk, debet, saldo. Attribute kredit dihilangkan karena tidak diperlukan datanya karena atribut kredit digunakan pada pembeli yang melakukan hutang tetapi di Gemilang Cell tidak diperkenankan membayar secara kredit. Setelah dilakukan data Cleaning maka bisa terlihat pada tabel 3.2

Tabel 3.2 Setelah dilakukan cleaning

NO	Id Transaksi	Tgl Transaksi	Keterangan	Debet	Saldo
1	11038695	08/05/2019 21:43:25	Simpati 50 ribu	49900.	375300.00
2	11038703	08/05/2019 21:46:55	Simpati 25 ribu	25350.0	349950.00
3	11040037	09/05/2019 12:27:42	Indosat 5 Ribu	6200.0	332550.00
4	11040039	09/05/2019 12:28:35	Indosat Sms 5 ribu	6500.0	326050.00
5	11043531	10/05/2019 16:41:11	Simpati 20 ribu	206500	940300.00
6	11043533	10/05/2019 16:41:37	Simpati 20 ribu	10800.0	929500.00
.					
.					
.					
.					
364	112666393	27/07/2019 00:16:08	Kuota Simpati 2GB	33650.00	1055900.00

3.3. Metode Association Rule

Proses aturan asosiasi yang akan dilakuakn pada penelitian ini dibagi menjadi 9 tahap sebagai berikut:

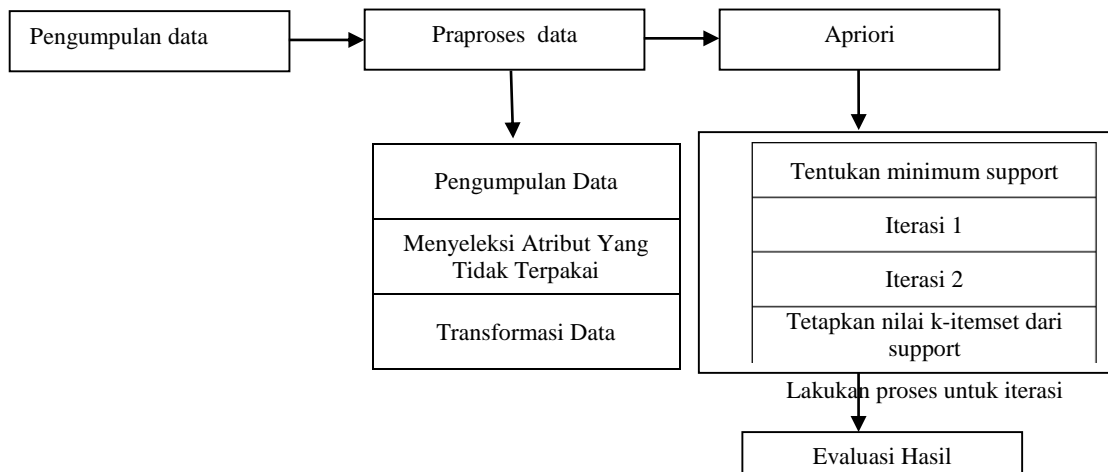
- 1 Sistem men-scan database untuk mendapatkan kandidat 1-itemset (himpunan item yang terdiri dari 1 item) dan menghitung nilai supportnya. Kemudian nilai supportnya tersebut dibandingkan dengan

minimum support yang telah ditentukan, jika nilainya lebih besar atau sama dengan minimum support maka itemset tersebut termasuk dalam large itemset.

2. Itemset yang tidak termasuk dalam large itemset tidak diikuti dalam iterasi selanjutnya (diprune).
3. Pada iterasi kedua sistem akan menggunakan hasil large itemset pada iterasi pertama (L_1) untuk membentuk kandidat itemset kedua (L_2). Pada iterasi selanjutnya sistem akan menggunakan hasil large itemset pada iterasi selanjutnya akan menggunakan hasil large itemset pada iterasi sebelumnya (L_{k-1}) untuk membentuk kandidat itemset berikut (L_k). Sistem akan menggabungkan (join) L_{k-1} dengan L_{k-1} untuk mendapatkan L_k , seperti pada iterasi sebelumnya sistem akan menghapus (prune) kombinasi itemset yang tidak termasuk dalam large itemset.
4. Setelah dilakukan operasi join, maka pasangan itemset baru hasil proses join tersebut dihitung supportnya.
5. Proses pembentuk kandidat yang terdiri dari proses join dan prune akan terus dilakukan hingga himpunan kandidat itemsetnya null, atau sudah tidak ada lagi kandidat yang akan dibentuk.
6. Setelah itu, dari hasil frequent itemset tersebut dibentuk association rule yang memenuhi nilai support dan confidence yang telah ditentukan.
7. Pada pembentukan association rule, nilai yang sama dianggap sebagai satu nilai.
8. *Association rule* yang terbentuk harus memenuhi nilai minimum yang telah ditentukan.
9. Untuk setiap large itemset L , kita cari himpunan bagian L yang tidak kosong. Untuk setiap himpunan bagian tersebut, dihasilkan rule dengan bentuk $aB(L-a)$ jika supportnya (L) dan supportnya (a) lebih besar dari minimum *support*.

3.4 Langkah Penelitian

Tahapan atau langkah dalam penelitian disajikan pada Gambar 3.1:



Gambar 3.1 Langkah Penelitian

Berikut merupakan langkah dalam melakukan penelitian yang dilakukan:

1. Pengumpulan data
2. Praproses data
3. Apriori
4. Langkah – langkah apriori
5. Evaluasi Hasil

3.5 Praproses Data

Tahapan atau Langkah Praproses Data Sebagai Berikut:

1. Pengumpulan Data

Pada tahapan ini sumber data yang digunakan berasal dari data transaksi penjualan pada *counter* Gemilang Cell. Data tersebut akan diproses guna menghasilkan pengetahuan yang bisa digunakan sebagai pengembangan strategi bisnis.

2. Menyeleksi atribut yang tidak terpakai

Pada tahapan ini akan dilakukan beberapa seleksi terlebih pada atribut yang sekiranya tidak diperlukan hasil dari menyeleksi atribut yang tidak terpakai seleksi ini dilakukan dengan cara data *cleansing* berikut data yang sudah dilakukan melalui cara *cleansing* dapat dilihat pada tabel 3.5

Tabel 3.5 Data Cleanig

NO	Id Transaksi	Tgl Transaksi	Keterangan	Debet	Saldo
1	111953571	01/07/2019 07:40:05	Simpati 25 ribu	25350.00	0.00
2	111953578	01/07/2019 07:40:20	Simpati 5 ribu	5800.00	0.00
3	111953671	01/07/2019 07:43:19	Simpati 50 ribu	49900.00	0.00
4	112062547	04/07/2019 17:31:19	Simpati 10	10900.00	0.00
5	112062561	04/07/2019 17:31:51	Simpati 5 ribu	5900.00	0.00
.					
.					
.					
.					
364	112666393	27/07/2019 00:16:08	Kuota Simpati 2GB	33650.00	1055900.00

3. Transformasi Data

Pada tahap ini dilakukan transformasi data sesuai dengan sistem yang akan digunakan dalam analisis data mining. Transformasi ini dilakukan dengan cara memasukkan data-data penggabungan dari pengumpulan data dan menyeleksi atribut tadi sehingga menjadi data yang sudah final. Maka didapatkan hasil setelah dilakukan transformasi data berikut dapat dilihat pada tabel 3.5

Tabel 3.5 Transformasi Data

NO	Id Transaksi	Tgl Transaksi	Keterangan	Debet	Saldo
1	11038695	08/05/2019 21:43:25	Simpati 50 ribu	49900.	375300.00

Lanjutan Dari Tabel 3.5 Transformasi Data

3	11040037	09/05/2019 12:27:42	Indosat 5 Ribu	6200.0	332550.00
4	11040039	09/05/2019 12:28:35	Indosat Sms 5 ribu	6500.0	326050.00
5	11043531	10/05/2019 16:41:11	Simpaty 20 ribu	206500	940300.00
6	11043533	10/05/2019 16:41:37	Simpaty 20 ribu	10800.0	929500.00
.					
.					
.					
.					
364	112666393	27/07/2019 00:16:08	Kuota Simpaty 2GB	33650.00	1055900.00

1.6 Algoritma Apriori

Untuk menentukan algoritma apriorinya maka dapat dilakukan dengan berbagai langkah sebagai berikut.

1. Tentukan minimum support dari data transaksi penjualan dari sampel data yang telah dimasukan .
2. Iterasi 1: hitung item-item dari support(transaksi yang memuat seluruh item) dengan men-scan database untuk 1-itemset, setelah 1-itemset didapatkan, dari 1-itemset apakah diatas minimum support, apabila telah memenuhi Sminimum support, 1-itemset tersebut akan menjadi pola frequent tinggi.
3. Iterasi 2: untuk mendapatkan 2-itemset, harus dilakukan kombinasi dari k-itemset sebelumnya, kemudian scan database lagi untuk hitung item-item yang memuat support. itemset yang memenuhi minimum support akan dipilih sebagai pola frequent tinggi dari kandidat.
4. Tetapkan nilai k-itemset dari support yang telah memenuhi minimum support dari k-itemset.
5. Lakukan proses untuk iterasi selanjutnya hingga tidak ada lagi k-itemset yang memenuhi minimum support.

Formula Pencarian Nilai Support & Confidence

Nilai support sebuah item diperoleh dengan rumus berikut:

$$SupportA = \frac{\text{jumlah Transaksi mengandung } A}{\text{Total Transaksi}} = \dots\dots(3.1)$$

Sedangkan nilai support dari 2 item diperoleh dari rumus berikut:

$$Support(A \cup B) = \frac{\text{jumlah Transaksi mengandung } A \text{ dan } B}{\text{Total Transaksi}} = \dots\dots(3.2)$$

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan assosiatif yang memenuhi syarat minimum untuk confidence dengan menghitung confidence aturan assosiatif A -> B. Nilai confidence dari aturan A -> B diperoleh dari rumus berikut:

$$Confidence = p\left(\frac{B}{A}\right) = \frac{\text{Jumlah Transaksi mengandung } A \text{ dan } B}{\text{Jumlah Transaksi mengandung } A} = \dots\dots(3.3)$$

3.4 Evaluasi Hasil

Model Asosiasi yang dibangun perlu di evaluasi untuk mengetahui seberapa baik model dalam melakukan klasifikasi dalam mengevaluasi kinerja Asosiasi dapat dilakukan dengan *software Weka*. Hasil dari implementasi model data mining akan terdapat sebuah *rule.rule* yang dihasilkan akan digunakan sebagai dasar prediksi *rule* yang akan dilakukan. Evaluasi dalam tahapan ini menggunakan Metode Association rule

