

BAB 5 PEMBAHASAN

Proses *data mining* merupakan proses berulang dan harus dilihat secara menyeluruh. Antara tahapan yang satu dengan selanjutnya harus dilalui secara konsisten dan berkelanjutan sampai ditemukan suatu pengetahuan baru yang berkaitan dengan permasalahan bisnis. Sukses tidaknya proses *data mining* ditentukan oleh tahapan-tahapan yang dikerjakan. Pada bab pembahasan ini penulis berusaha menjelaskan jawaban dari pertanyaan-pertanyaan dan pencapaian dari tujuan penelitian ini.

5.1 Pertanyaan dan Tujuan dari Penelitian

5.1.1 Pertanyaan Pertama

Bagaimana penerapan teknologi *data mining* untuk memroses data *billing/usage* pelanggan *wireline* dan kinerja *network* ?

Pembahasan:

Dengan terlebih dahulu memahami permasalahan bisnis yang ada termasuk kebutuhan perusahaan terhadap *data mining*, maka data *usage* dan kinerja *network* dikumpulkan dalam suatu *data warehouse*. Kemudian dilanjutkan dengan proses *data mining* dengan metodologi *Data Mining Two Crows-Corporation*. Adapun pertimbangan menggunakan metodologi ini adalah tahapan-tahapan yang harus dilalui cukup lengkap dan terintegrasi mulai dari tahapan *understanding bussiness problem* (pemahaman permasalahan bisnis) sampai dengan tahapan *Deploy model and result* (menjalankan model dan bagaimana hasilnya). Di samping itu metodologi tersebut mereferensi metedologi *CRISP-DM*.

5.1.2 Pertanyaan Kedua

Bagaimana melakukan analisa data hasil OLAP untuk digunakan dalam membantu proses *data mining*?

Pembahasan:

Dengan adanya Teknologi OLAP membantu mengeksplorasi data data yang sudah terkumpul dalam satu *data warehouse*. Dari data dapat dianalisa secara periode misalnya harian, mingguan, bulanan, tahunan dan seterusnya. Kemudian dapat dilakukan analisa pendalaman data misalnya dari area bisnis data dapat di dalam ke area lebih kecil yaitu area *switching* dan seterusnya. Serta dengan menggunakan OLAP dapat diketahui pola-pola yang dominan dari suatu data. Misalnya dalam penelitian ini *usage* lokal dan *PONSEL* merupakan *usage* dominan bila dilihat secara periode maupun di semua area *switching*. Hasil OLAP merupakan pengetahuan awal terhadap pola-pola data yang ada yang diharapkan dapat membantu mencari pengetahuan baru lainnya yang lebih spesifik di tahapan selanjutnya.

5.1.3 Pertanyaan Ketiga

Bagaimana mencari pengetahuan mengenai pola-pola *service usage* dan kinerja *network*?

Pembahasan:

Mulai tahapan *data explore* yaitu OLAP sudah dapat diperoleh pengetahuan awal terhadap pola-pola data yang ada yaitu *usage* dominan secara berurutan adalah ditemukan fakta bahwa komposisi *Usage* layanan mulai yang terbesar sampai dengan yang terkecil adalah : *PONSEL* (1.65%-3.75%), *Lokal* (2.29%-2.91%), *SLJJ* (1.07%-1.41%), *INET* (0.40%-0.52%), *SLI007*(0.33%-0.44%), *SLI001*(0.02%-0.03%) dan *SLI008*(0.001%-0.002%).

Kemudian dilakukan *data mining* dengan menggunakan model pencarian *Association Rules* yaitu *algoritma apriori* terhadap data *usage*, *fault* dan *tenure*.

Adapun pengetahuan baru yang dihasilkan:

1. Pola *cross selling*/layanan dengan *usage* rendah namun dominan.

- *Association Rules*: Lok26rb → Sel40rb (49%,71%)

Hal ini berarti bahwa *customer* yang menggunakan *usage* local rata-rata Rp 26.000,- dan seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian dari *customer* menggunakan *usage* lokal rata-rata sebanyak

Rp 26.000 kemungkinan menggunakan seluler rata-rata Rp 40.000, sebesar 71%.

- *Association Rules: Sel40rb → Lok26rb (51%,69%)*.

Hal ini berarti bahwa *customer* yang menggunakan *usage* seluler rata-rata Rp 40.000,- dan lokal rata-rata Rp 26.000 sebanyak 51%. Kemudian *customer* yang menggunakan seluler rata-rata Rp 40.000 kemungkinan menggunakan lokal rata-rata Rp 26.000, sebesar 69%.

- *Association Rules: Sel40rb → Sljj31rb (28%,62%)*

Hal ini berarti bahwa *customer* yang menggunakan *usage* seluler rata-rata Rp 40.000,- dan SLJJ rata-rata Rp 31.000 sebanyak 28%. Kemudian *customer* yang menggunakan seluler rata-rata Rp 31.000,- kemungkinan menggunakan SLJJ rata-rata Rp 31.000,- sebesar 62%.

- *Association Rules: Sljj31rb → Sel40rb (49%,36%)*

Hal ini berarti bahwa *customer* yang menggunakan *usage* SLJJ rata-rata Rp 31.000,- dan seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian *customer* yang menggunakan SLJJ rata-rata Rp 31.000,- kemungkinan menggunakan seluler rata-rata Rp 40.000, sebesar 36%

2. Adanya *customer* yang sudah lama berlangganan, namun dengan pola pemakaian/ *usage* yang masih rendah.

- *Association Rules: Ten13thn → Lok26rb (50%,34%)*

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 13 tahun dan menggunakan *usage* lokal rata-rata Rp 26.000 sebanyak 50%. Kemudian yang sudah berlangganan 13 tahun dengan kemungkinan pemakaian lokal Rp 26.000 sebesar 34%.

- *Association Rules: Ten13thn → Sel40rb (49%,32%)*

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 13 tahun dan menggunakan Seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian yang sudah berlangganan 13 tahun dengan kemungkinan pemakaian seluler Rp 40.000 sebanyak 32%.

- *Association Rules*: Ten26thn → lok26rb (50%,30%)

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 26 tahun dan menggunakan lokal rata-rata Rp 26.000 sebanyak 50%. Kemudian yang sudah berlangganan 26 tahun dengan kemungkinan pemakaian lokal Rp 26.000, sebesar 30%.

- *Association Rules*: Ten26thn → Sel40rb (49%,31%)

Hal ini berarti bahwa *customer* yang sudah berlangganan rata-rata 26 tahun dan menggunakan seluler rata-rata Rp 40.000 sebanyak 49%. Kemudian yang sudah berlangganan 26 tahun dengan kemungkinan pemakaian seluler Rp 40.000, sebesar 31%.

3. *Fault duration* yang lama dapat berkontribusi terhadap kecilnya kemungkinan *customer* menggunakan layanan.

- *Association Rules*: FD11JAM → SEL40RB (49%,1%)

Hal ini berarti *customer* yang mengalami *Fault duration* rata-rata 11 jam dan pemakaian *usage* seluler rata-rata Rp 40.000 sebanyak 49%, kemudian yang mengalami *fault duration* rata-rata 11 jam kemungkinan menggunakan seluler rata-rata Rp 40.000, hanya sebesar 1%.

- *Association Rules*: FD11JAM → SLJJ31RB (28%,1%)

Hal ini berarti *customer* yang mengalami *Fault duration* rata-rata 11 jam dan pemakaian *usage* SLJJ rata-rata Rp 31.000 sebanyak 28%, kemudian yang mengalami *fault duration* rata-rata 11 jam kemungkinan menggunakan seluler rata-rata Rp 40.000, hanya sebesar 1%.

- *Association Rules*: FD11JAM → LOK142RB (24%,1%)

Hal ini berarti *customer* yang mengalami *Fault duration* rata-rata 11 jam dan pemakaian *usage* Lokal rata-rata Rp 142.000 sebanyak 24%, kemudian yang mengalami *fault duration* rata-rata 11 jam kemungkinan menggunakan seluler rata-rata Rp 142.000, hanya sebesar 1%.

5.1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk *Discovery Knowledge* data *billing/usage* dan kinerja *network* pelanggan *wireline*.

Pembahasan:

Dari penelitian ini menghasilkan pengetahuan baru berkaitan dengan data *usage billing* maupun data kinerja *network*. Pola-pola pengetahuan tersebut muncul ketika dilakukan *explore data* dengan teknologi OLAP dan dengan menggunakan model pencarian pengetahuan dengan Association Rules Algoritma *apriori*. Dalam proses OLAP diperoleh pengetahuan awal, setelah itu dilakukan proses pencarian dengan algoritma *apriori* diperoleh pengetahuan baru lainnya yang lebih berarti. Dengan diperolehnya pengetahuan-pengetahuan baru melalui tahapan-tahapan tersebut, maka dapat dikatakan bahwa tujuan dari penelitian ini tercapai.

