

DATA WAREHOUSE, DATA MINING DAN KONSEP CROSS-SELLING PADA ANALISIS PENJUALAN PRODUK

Eka Miranda; Natalya Elfreida

Jurusan Sistem Infromasi, Fakultas Ilmu Komputer, Bina Nusantara University
Jln. K.H. Syahdan No.9, Palmerah, Jakarta Barat 11480
ekamiranda@yahoo.com, natalya_3a2@yahoo.co.id

ABSTRACT

This paper is about designing and implementing data warehousing and data mining, along with their roles in supporting decision-making related to sales product analysis in cross-selling concept of PT XYZ. The database the company used is not supporting data analysis and decision-making. Therefore, it made a data warehousing design that could be used to keep data in a huge amount and could give report and answer from user's questions in ad hoc. The method is used to design and implement data warehousing and data mining which consists of literature study, company problem analysis, and data warehousing design, and testing result. The writing results are a data warehousing design and data mining and also the implementation of cross-selling concept to analysis the sales, purchases, and customers' cancellation data. The data could be showed and analyzed from some point of views that could help managers to analyse and acknowledge more information.

Keywords: *decision, cross-selling, data warehousing, data mining.*

ABSTRAK

Tulisan ini membahas mengenai rancangan dan implementasi data warehouse dan data mining serta peranannya dalam mendukung pengambilan keputusan yang terkait dengan analisis penjualan produk dengan konsep cross-selling pada PT XYZ. Database yang digunakan perusahaan saat ini tidak mendukung analisa data dan pengambilan keputusan. Berdasarkan permasalahan ini maka dibuat desain data warehouse yang dapat digunakan untuk menyimpan data dalam jumlah besar dan memungkinkan untuk memberikan pelaporan dan jawaban dari pertanyaan pengguna yang bersifat ad hoc. Metode yang digunakan untuk desain dan implementasi data warehouse dan data mining terdiri dari studi literatur, analisis permasalahan perusahaan, perancangan data warehouse, dan pengujian hasil. Hasil dari tulisan ini berupa desain data warehouse dan data mining serta implementasinya dengan konsep cross-selling untuk menganalisis data penjualan, pembelian dan pembatalan oleh pelanggan. Data tersebut dapat ditampilkan dan dianalisis dari beberapa sudut pandang yang berbeda, yang dapat membantu manager untuk menganalisis dan menggali lebih banyak lagi informasi.

Kata kunci: *keputusan, cross-selling, data warehouse, data mining*

PENDAHULUAN

Latar Belakang

Dalam pengelolaan basis data (*database*) organisasi baik organisasi bisnis maupun organisasi non profit terdapat istilah bahwa organisasi kaya akan data, tetapi miskin akan informasi atau *businesses have been rich in data and poor in information*. Data warehouse bukan hanya suatu arsip atau fasilitas penyimpanan data, *data warehouse* harus dengan cerdas dirancang untuk berisi informasi dan data yang penting untuk mendukung memecahkan masalah secara aktif. *Data warehouse* dikembangkan untuk mempertemukan suatu permintaan bagi Manajemen Informasi dan Analisa yang tidak bisa dijumpai pada basis data. Aplikasi basis data tidak mampu untuk mempertemukan kebutuhan ini untuk beberapa pertimbangan: Pengolahan pelaporan data yang banyak/besar dapat mengurangi waktu tanggap aplikasi basis data, Perancangan aplikasi basis data tidaklah dioptimalkan untuk analisa informasi dan pelaporan, Kebanyakan organisasi mempunyai lebih dari satu aplikasi basis data, maka pelaporan secara keseluruhan organisasi tidak bisa didukung dari sistem tunggal, pengembangan laporan pada aplikasi basis data sering memerlukan program komputer spesifik yang mahal.

Selain itu hal yang penting dihadapi oleh perusahaan saat ini adalah persaingan. Karena itu diperlukan pemasaran yang berorientasi kepada pelanggan dan dibutuhkan informasi yang memanfaatkan *database* yang menyimpan transaksi pelanggan untuk melakukan analisis penjualan produk atau jasa dengan teknik data mining dan konsep *cross-selling*. Di sini perlu mengkombinasikan data hasil survei dengan data transaksi pelanggan yang disimpan di dalam *database* untuk kemudian dihitung probabilitas dan dilakukan prediksi penjualan produk dan jasa perusahaan (Kamakura, Wedel, Rosa, & Mazzon, 2003). Data warehouse sebagai *database* yang khusus dirancang untuk mendukung Manajemen Informasi dan Analisa perlu untuk dikembangkan. *Data warehouse* dapat mengintegrasikan data dari berbagai sumber data yang berbeda dan mengintegrasikan informasi ini di dalam tempat tunggal.

Perumusan Masalah

Masalah yang dapat dirumuskan pada analisis penjualan produk PT. XYZ terdiri atas bagaimana menghasilkan informasi yang dibutuhkan oleh manager perusahaan mengenai fakta-fakta bisnis seperti pembelian, penjualan dan pembatalan pembelian oleh pelanggan yang terjadi pada suatu periode? Bagaimana memproses data menjadi informasi yang membantu para manager mengambil keputusan? Misalkan laporan penjualan serta produk apa yang paling laku dan informasi lainnya, Bagaimana membantu para manager dalam mengambil keputusan yang cepat dan tepat? Dari ketiga hal tersebut, dapat disimpulkan bahwa dibutuhkan tidak hanya suatu tempat penyimpanan seluruh data transaksi yang terjadi tapi dapat juga menghasilkan laporan yang berguna bagi pemilik perusahaan sebagai bahan pertimbangan untuk menganalisis suatu masalah dan mengambil keputusan.

Tujuan Penelitian

Tujuan dari penulisan hasil penelitian ini adalah untuk mempelajari dan menjelaskan data warehouse, data mining dan penerapannya untuk pengembangan bisnis dalam hal ini analisis terhadap penjualan produk perusahaan dan peningkatan kinerja perusahaan. Diharapkan Data Warehouse dan *data mining* dapat membantu perusahaan dalam meningkatkan efektifitas kerja dengan cara mempermudah analisis suatu masalah dan membantu dalam membuat keputusan yang cepat serta tepat berdasarkan informasi yang tersedia.

Studi Pustaka

Data Warehouse

Membuat *data warehouse* tidak sekedar memindahkan data operasional ke dalam *data warehouse*, memindahkan data dari beberapa sumber berpotensi menimbulkan masalah besar terutama terkait dengan isu tidak konsistennya data. *Data warehouse* tidak hanya digunakan dalam melakukan *loading*, integrasi, dan menyimpan data dalam jumlah besar, namun juga berpotensi untuk mendapatkan sudut pandang baru dari sebaran data, dan memungkinkan untuk memberikan pelaporan dan jawaban dari pertanyaan pengguna yang bersifat *ad hoc* secara cepat dan lebih baik. Tahapan untuk membangun data warehouse terdiri atas identifikasi sumber data, membangun ETL, ekstraksi, transformasi, dan *loading*.

Identifikasi sumber data

Langkah pertama sebelum mulai mengembangkan data warehouse yaitu identifikasi sumber data. Perlu mencari tahu apa data yang dibutuhkan untuk diletakkan ke dalam data warehouse.

Membangun ETL (*Extraction, Transformation, Loading*)

Setiap *data warehouse* memiliki data yang berbeda persyaratan, oleh karena itu, *ETL tool* yang sesuai adalah solusi yang lebih baik untuk memenuhi persyaratan.

Ekstraksi

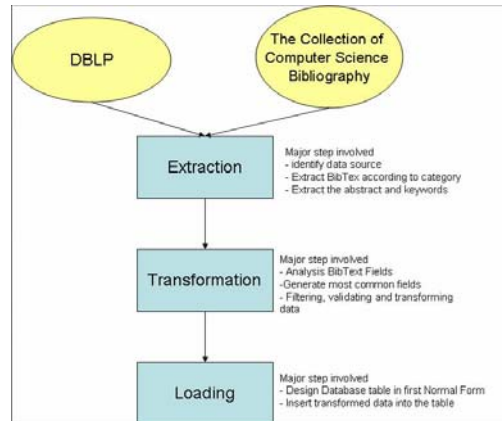
Perlu menentukan sistem *database* yang akan digunakan dan juga mencari tahu data apa yang diperlukan sebelum mendapatkannya. Penurunan biaya perangkat keras dan penyimpanan telah mengatasi masalah pada menghindari duplikasi data dan juga kekhawatiran pada kurangnya ruang penyimpanan sebagai tempat penyimpanan data yang berlebihan atau yang tidak diperlukan. Namun, mungkin tidak ada alasan untuk menyimpan data yang tidak diperlukan dan telah diidentifikasi tidak berguna dalam proses pengambilan keputusan. Oleh karena itu, diperlukan untuk mendapatkan hanya ekstrak data yang relevan sebelum membawa ke *data warehouse* (Mallach, 2000).

Transformasi

Setelah penggalian data dari berbagai sumber, transformasi dibutuhkan untuk menjamin konsistensi data. Agar mengubah data ke dalam data warehouse dengan benar, perlu mengetahui cara pemetaan bidang sumber data eksternal ke data warehouse. Transformasi dapat dilakukan selama ekstraksi data atau saat memuat data ke dalam data warehouse. Integrasi ini bisa menjadi masalah yang kompleks ketika jumlah sumber data menjadi lebih besar.

Loading

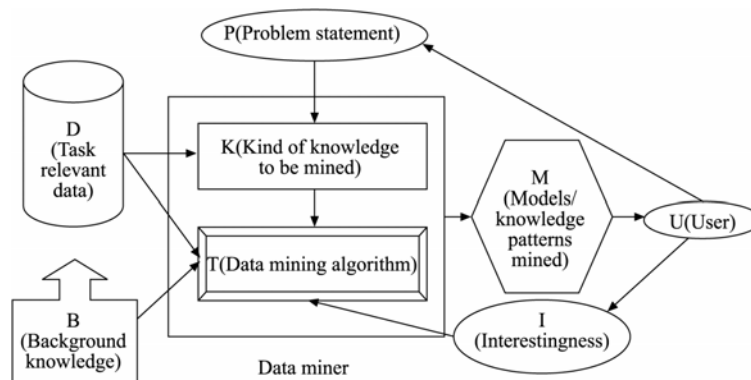
Setelah proses penggalian, mengubah dan pembersihan telah dilakukan, data diambil ke dalam data warehouse. Pemuatan data dapat dikategorikan ke dalam dua jenis; pemuatan data yang saat ini ada dalam *database* operasional dan pemuatan pembaruan *data warehouse* dari perubahan yang telah terjadi dalam *database* operasional. Untuk menjamin kesegaran data, *data warehouse* perlu untuk terus memperbaharui datanya. Banyak persoalan yang perlu untuk dipertimbangkan terutama saat memuat data yang sudah diperbaharui ke data warehouse. Sementara melakukan pemutakhiran data warehouse, perlu memastikan bahwa tidak ada data yang dilepaskan dan juga untuk memastikan *overhead* minimum selama proses *scanning* file yang ada.



Gambar 1. Tahap-tahap membangun data warehouse
 Sumber: Wah, Peng & Hok (2007)

Data Mining

Data mining (DM) adalah proses mencari dan menggali hubungan antar data yang berguna bagi penggunaannya. *Data mining* (DM) juga merupakan proses mencari data untuk menemukan yang sebelumnya tidak diketahui dari hubungan antara data yang menarik untuk penggunaannya (Hand, 1998). Tahap terpenting dalam mendapatkan pengetahuan dari dalam *database* adalah *data mining*. DM telah menjadi bidang ilmu tersendiri (Fayyad, Piatetsky-Shapiro & Smyth, 1996; Chen & Liu, 2005). *Data mining* adalah subjek yang terus berkembang (Chen & Liu, 2005) di mana digunakan untuk mendapatkan pengetahuan (*knowledge discovery*) dari dalam *database* yang mengacu pada keseluruhan proses mulai dari *low-level* data sampai ke *high-level* data. *Data mining* adalah proses untuk menemukan suatu pola dan *trend* dari data. Tujuan dari proses ini adalah menyusun data yang berjumlah besar untuk mendapatkan informasi baru. *Data mining* juga dapat berarti eksplorasi dan analisis data secara otomatis atau semi otomatis dari data yang jumlahnya besar untuk mendapatkan pola dan aturan yang bermakna bagi penggunaannya (Fayyad, Piatetsky-Shapiro & Smyth, 1996). Chang & Chen (2006) dan Yu & Chen (2001) menjelaskan bahwa proses data mining terdiri atas enam proses kerja utama yang dapat dilihat pada Gambar 2.



Gambar 2. Proses kerja data mining
 Sumber : Chang & Chen (2006)

Data mining adalah proses mencari tren dan pola dalam data. Tujuan dari proses ini adalah untuk menyortir besar jumlah data dan menemukan informasi baru. Manfaat dari *data mining* adalah

pengetahuan baru ditemukan dan hasil ditindaklanjuti. Tujuan dari data mining adalah untuk mengekstrak informasi yang berharga dan baru dari data yang ada. Secara umum, data mining meliputi fungsi utama, yaitu: Klasifikasi, Clustering, Estimasi, Prediksi, dan Deskripsi.

Teknologi *data mining* dapat dibagi antara teknologi tradisional dan halus. Statistik analisis merupakan perwakilan dari teknologi tradisional. Adapun teknologi *data mining* halus adalah semua jenis kecerdasan buatan yang digunakan, yang lebih sering digunakan adalah jenis pohon keputusan, jaringan syaraf, algoritma genetika, logika *fuzzy* dan aturan induksi. Penggunaan berbagai jenis aplikasi yang berbeda sering dapat menyebabkan hasil yang berbeda.

Han & Kamber (2001), dan Buja & Lee (2001), model *data mining* yang umum digunakan adalah (1) *classification*, mengklasifikasikan item data ke dalam salah satu dari beberapa standar kategoris kelas; (2) *regression*, memetakan sebuah item data ke sebuah variabel nilai prediksi nyata; (3) *clustering*, memetakan sebuah item data ke dalam salah satu dari beberapa *cluster*, dimana *cluster* mengelompokkan item data berdasarkan matrik kesamaan atau model kepadatan probabilitas; (4) *rule generation*, ekstrak aturan klasifikasi dari data; (5) *discovering association rule*, asosiasi menggambarkan hubungan antara berbagai atribut; (6) *summarization*, memberikan gambaran solid untuk subset data; (7) *dependency modelling*, menggambarkan dependensi yang signifikan antara variabel.

Star Schema

Star schema merupakan paradigma modeling yang paling banyak digunakan, di mana di dalamnya mengandung antara lain sebuah tabel pusat yang besar tanpa adanya data *redundancy* di dalamnya, yang biasa disebut dengan tabel fakta. Selain itu, di dalam *star schema* juga mengandung satu set tabel yang lebih kecil, yang biasa disebut dengan tabel dimensi. Berikut adalah beberapa karakteristik utama dari *star schema* antara lain: pusat dari *star schema* adalah tabel fakta, tabel fakta berisi indikator-indikator kinerja pokok, objek-objek informasi dan waktu adalah kunci utama tabel fakta, tabel-tabel yang ada di sekeliling tabel fakta adalah tabel dimensi, tabel dimensi berisi data mengenai objek-objek informasi atau waktu, tabel fakta dan tabel dimensi direlasikan dengan *key* yang ada, *Star scheme* diimplementasikan menggunakan teknologi *relational database*.

Cross-Selling

Sebuah aspek penting dari orientasi baru pada pemasaran berorientasi pelanggan adalah penggunaan *database* transaksi pelanggan untuk analisis menjual-silang (*cross-selling*) layanan dan produk. *Cross-selling* berkaitan dengan upaya peningkatan jumlah produk atau jasa yang pelanggan gunakan atau beli dari perusahaan. *Cross-selling* produk dan jasa yang dibeli pelanggan saat ini memerlukan biaya lebih rendah dari biaya yang terkait untuk memperoleh pelanggan baru, karena perusahaan sudah memiliki hubungan dengan pelanggan. Implementasi yang tepat dari pelaksanaan *cross-selling* hanya dapat dicapai jika ada suatu infrastruktur informasi yang memungkinkan manajer untuk menawarkan produk atau layanan kepada pelanggan yang menyerap kebutuhan mereka tetapi belum dijual kepada mereka (Kamakura, Wedel, Rosa & Mazzon, 2003). Pada prinsipnya, strategi pemasaran *cross-selling* adalah sebuah istilah umum yang digunakan untuk menjelaskan penjualan tambahan produk atau layanan kepada pelanggan yang telah membeli sesuatu dari perusahaan. Strategi ini berhubungan dengan analisis data pelanggan.

Dengan menganalisa perilaku pelanggan, dapat dimungkinkan untuk memperkirakan model perilaku pelanggan selanjutnya dan juga perilaku pelanggan mana yang potensial. Hal ini sangat menguntungkan bagi perusahaan karena dapat menjual produk dan layanan yang berbeda dengan menggunakan teknik pemasaran yang lebih beragam melalui pengetahuan (*knowledge*) yang telah diperolehnya. Perusahaan dapat mengidentifikasi kebutuhan dan keinginan pelanggan untuk

menetapkan *additional product* dari produk utamanya. *Additional product* yang disarankan dapat berupa *common-sense-based* (contohnya, jika pelanggan membeli sebuah kamera digital, maka *cross-sell product*-nya biasanya adalah *memory card*, *case*, dan *spare-battery*) atau dapat pula berupa *data-driven* (*cross-selling* didapat dari data pembelian historikal pelanggan). *Cross-selling* biasanya digunakan oleh sebagian besar *online store* untuk menentukan rekomendasi produk apa yang seharusnya dijual juga. Sebagai contoh dapat dilihat salah satu *online bookstore* terkemuka seperti Amazon.com dapat dengan jelas terlihat bahwa jika seorang pelanggan membeli buku secara *online* maka *website* akan memberikan pula rekomendasi mengenai buku yang terkait yang direkomendasikan untuk dibeli. Hal ini dapat dilakukan melalui analisis.

Database Pemasaran

Database pemasaran melibatkan pembangunan, pengorganisasian, melengkapi, dan menggali transaksi pelanggan untuk meningkatkan akurasi dari upaya pemasaran dan memungkinkan identifikasi prospek terbaik untuk upaya pemasaran (Goodman, 1992).

METODE

Tahapan yang dilakukan untuk mengembangkan dan implementasi *data warehouse*, teknik *data mining* dan *cross-selling* dalam penelitian ini terdiri dari (1) studi literatur, mengumpulkan dan mempelajari bahan-bahan yang diperlukan untuk pembuatan *data warehouse*; (2) analisis perusahaan, mengumpulkan data dari perusahaan yang dibutuhkan dalam pembuatan *data warehouse* dengan melakukan wawancara dan pengamatan secara langsung ke perusahaan dan melakukan analisis terhadap sistem basis data yang dijalankan oleh perusahaan; (3) perancangan *data warehouse*, membuat Rancangan *Entity Relationship Diagram* (ERD), Menetapkan tabel tabel yang digunakan dalam *data warehouse* dan Membuat rancangan *Star Schema*; (4) pengujian hasil, melakukan pengujian terhadap rancangan *data warehouse* dan *data mining* telah selesai dibuat sehingga dapat diketahui apakah telah berjalan sesuai dengan harapan dan tujuan yang ingin dicapai; (5) kesimpulan, setelah pengujian sudah selesai dilakukan dan diperoleh hasil yang diinginkan, maka hasil dapat diterapkan.

HASIL DAN PEMBAHASAN

Analisis Kebutuhan

Pada fase ini mendefinisikan semua sumber data yang digunakan dalam *data warehouse* dan mendefinisikan kebutuhan informasi yang harus dihasilkan oleh *data warehouse* melalui analisis *reporting*.

Kebutuhan Sumber Data

Sumber data yang digunakan dalam *data warehouse* adalah data yang ada dalam *database* operasional yang terdiri atas tabel MsBuku, MsCabang, MsAksesoris, MsJenisBuku, MsJenisKaryawan, MsJilidBuku, MsKaryawan, MsPelanggan, MsProvinsi, MsSupplier, DetailPembelianAksesoris, DetailPembelianBuku, DetailPenjilidan, DetailPenjualanAksesoris, DetailPenjualanBuku, HeaderPembelianAksesoris, HeaderPembelianBuku, HeaderPenjilidan, Tabel HeaderPenjualanAksesoris, dan Tabel HeaderPenjualanBuku.

Kebutuhan Informasi

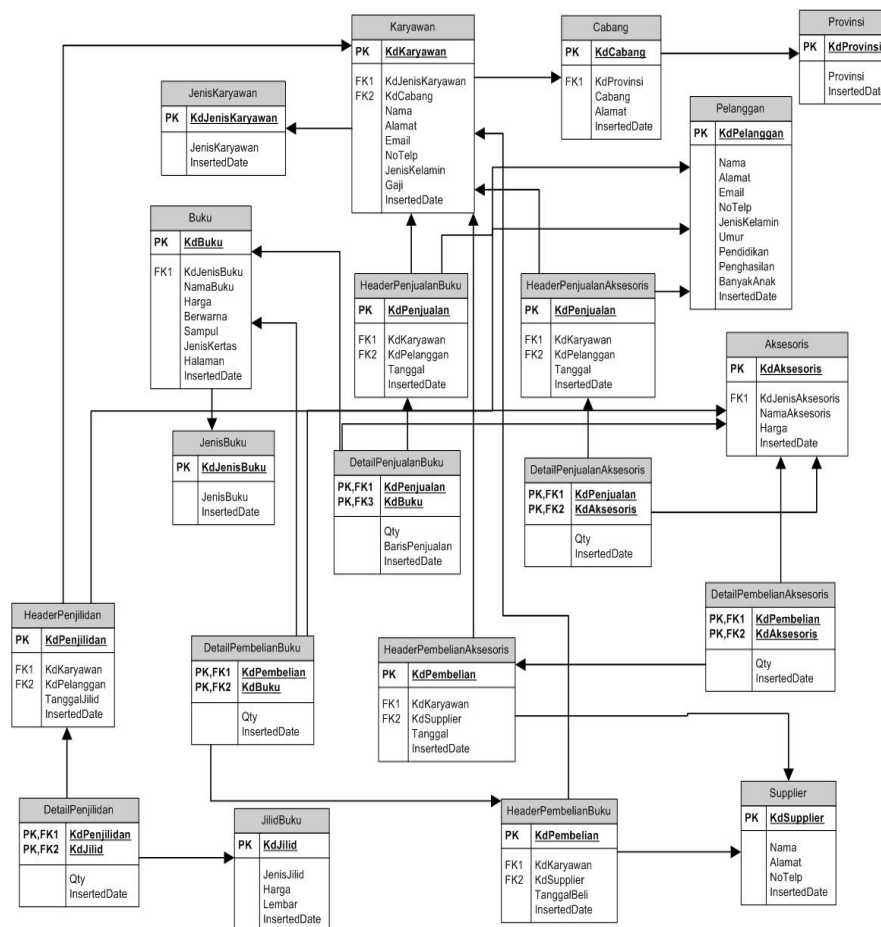
Informasi yang dibutuhkan oleh pihak manajemen untuk mendukung analisis dalam mengambil keputusan terdiri atas tren terhadap buku yang dibeli oleh pelanggan, Informasi mengenai jenis buku yang sering dibeli, tren proses pembelian oleh pelanggan, tren terhadap aksesoris yang dibeli pelanggan, dan tren terhadap pembelian buku dan aksesoris yang dibeli oleh pelanggan.

Perancangan Data Warehouse

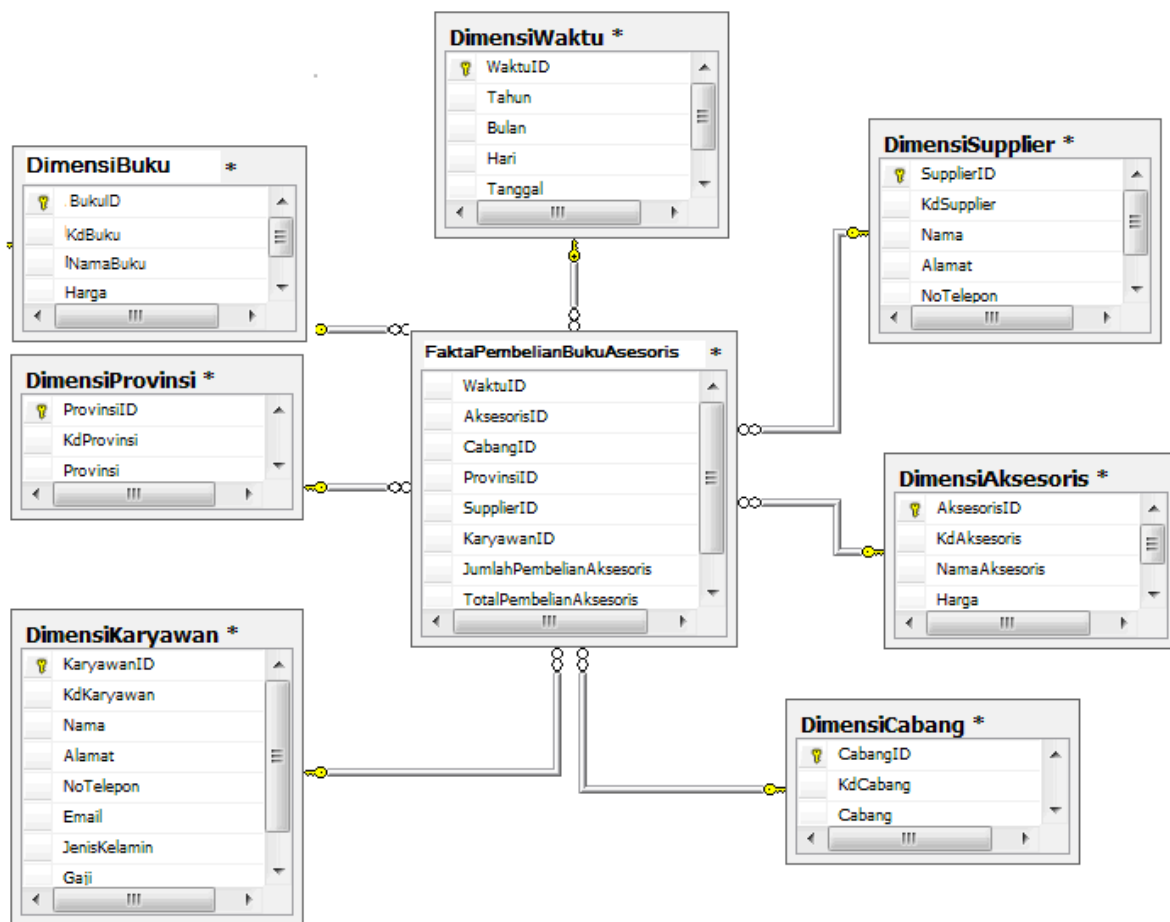
Data yang digunakan untuk membangun *data warehouse* berasal dari *database* operasional. Adapun gambar *Entity Relationship Diagram* (ERD) dari *database* yang digunakan dapat dilihat pada Gambar 3 dibawah ini:

Model Data Dimensional

Berdasarkan *Entity Relationship Diagram* (ERD) dan kebutuhan informasi yang dibutuhkan oleh manajemen maka dibuat model data dimensional dalam bentuk *star schema* untuk *data warehouse*. *Star schema* tersebut terdiri atas tabel dimensi, yaitu dimensi waktu, dimensi supplier, dimensi provinsi, dimensi aksesoris, dimensi cabang dan dimensi karyawan; dan tabel fakta yaitu fakta pembelian buku dan aksesoris.



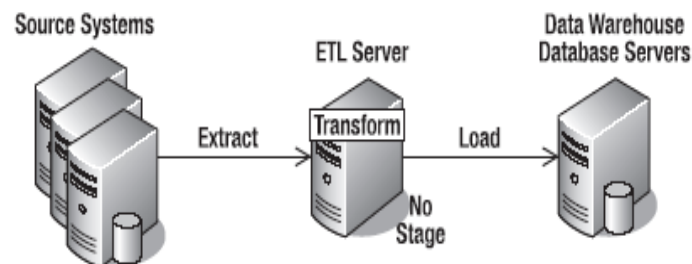
Gambar 3. Entity Relationship Diagram (ERD)



Gambar 4. Star Schema untuk data warehouse

Desain Proses ETL

Proses selanjutnya setelah *star schema* dibuat adalah membuat desain proses ETL (*Extraction, Transformation* dan *Loading*). Desain proses ETL, seperti pada Gambar 5, yang dipakai bisa dilihat pada gambar dibawah ini. Proses ETL mengambil data dari *source systems* menggunakan *query*. Setelah data hasil *query* diambil langkah selanjutnya dilakukan eksekusi proses ETL dan mengirimnya ke database *data warehouse*.



Gambar 5 Desain proses ETL

Pengujian Hasil

Penentuan Cross-Selling

Pemilihan produk didasarkan pada frekuensi (seberapa sering) kemunculan produk selama periode transaksi berlangsung. Dari data frekuensi tersebut, kemudian akan ditentukan tiga indikator yang berbeda untuk menghasilkan kemungkinan *cross-selling* yang paling baik. Tiga indikator tersebut adalah *support score*, *confidence score*, dan *improvement score*. *Support score* mengindikasikan seberapa sering dua produk muncul bersama-sama di dalam suatu periode transaksi, *confidence score* mengindikasikan seberapa banyak sebuah produk tertentu tergantung dengan produk yang lain, dan *improvement score* menunjukkan seberapa besar kemungkinan pelanggan yang lebih menyukai untuk membeli produk kedua (dalam hal ini *cross-selling product*) yang diberikan pada saat membeli produk pertama (*base product*).

Kemungkinan untuk menghasilkan *cross-selling product* yang terbaik dari total transaksi per pelanggan pada periode tertentu adalah item dengan *confidence score* yang tinggi. Jadi misalkan produk pertama adalah kombinasi A dan B dengan *confidence score* x% berarti bahwa ketika produk A dibeli, kemungkinan produk B juga dibeli sebesar x% dari seluruh periode transaksi.

Indikator kedua yang harus ditentukan adalah *support score* yang menentukan seberapa banyak produk A dan produk B muncul bersama-sama di dalam seluruh transaksi. Misalkan kombinasi produk C dan D memiliki *support score* y% mengindikasikan bahwa produk C dan produk D muncul secara bersama-sama sebesar y% dari seluruh total transaksi, sedangkan untuk kombinasi produk E dan F misalnya memiliki *improvement score* ≥ 1 , mengindikasikan bahwa produk E dan F merupakan *positive correlated*, yang berarti bahwa jika pelanggan membeli produk E, pelanggan juga setuju untuk membeli produk F, sedangkan jika *improvement score* < 1 , produk E dan F merupakan *negative correlated*. Produk dengan *negative correlated* tidak dapat dijadikan sebagai saran dalam penetapan *cross-selling product* walaupun memiliki *support score* dan *confidence score* yang tinggi.

Setelah dilakukan proses pencarian asosiasi antara produk yang satu dengan produk yang lain dari *dataset* yang ada maka diperoleh sekumpulan *rules* dan *itemset* seperti tampak pada Gambar 6. Dengan mengacu kepada pemodelan dalam menentukan *cross-selling product* di atas, maka dari seluruh kombinasi yang ada, seluruh *rules* yang mengandung kombinasi dua produk akan diteliti, namun hanya kombinasi yang mengandung unsur masing-masing produk yang menduduki peringkat 20 teratas sebagai produk pertama dari kombinasi tersebut yang akan digunakan selanjutnya.

Sesuai dengan kriteria penetapan *cross-selling product* yang terbaik, maka *improvement score* ditetapkan harus lebih besar atau sama dengan satu, karena *improvement score* kurang dari satu mengindikasikan pelanggan yang membeli *base product* tidak setuju untuk membeli *cross-selling product* yang disarankan, dan saran *cross-selling product* untuk masing-masing 20 produk tersebut akan dipilih dan diurutkan berdasarkan *confidence score*-nya. Dari tabel tersebut *confidence score* dan *improvement score*-nya masing-masing dinyatakan dalam persen, sedangkan *support score* dinyatakan dengan nilai yang mengindikasikan jumlah transaksi antara produk pertama (*base product*) dan *cross-selling product* dari total seluruh transaksi yang ada.

Y	Support	Size	Itemset
167	2		FILM WARNA NEW SUPERIA X-TRA 400/36, FILM WARNA NEW SUPERIA 200/36
146	2		FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA 200/36
123	2		AMPLOP FIS, FILM WARNA NEW SUPERIA 200/36
120	2		FILM WARNA NEW SUPERIA 200/20, FILM WARNA NEW SUPERIA 200/36
117	2		FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA X-TRA 400/36
115	2		FILM WARNA HORI COLOR GX 100 135/36, FILM WARNA NEW SUPERIA 200/36
112	3		FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA X-TRA 400/36, FILM WARNA NEW SUPERIA 200/36
101	2		FILM WARNA NEW SUPERIA 200/20, FILM WARNA NEW SUPERIA X-TRA 400/36
98	2		AMPLOP FIS, FILM WARNA NEW SUPERIA X-TRA 400/36
97	3		AMPLOP FIS, FILM WARNA NEW SUPERIA X-TRA 400/36, FILM WARNA NEW SUPERIA 200/36
94	3		FILM WARNA NEW SUPERIA 200/20, FILM WARNA NEW SUPERIA X-TRA 400/36, FILM WARNA NEW SUPERIA 200/36
93	2		FILM WARNA PRO-PLUS (3 ROLL), FILM WARNA NEW SUPERIA 200/36
92	2		AMPLOP FIS, FILM FUJICOLOR 200
91	2		FILM WARNA NEW SUPERIA 200/20, FILM FUJICOLOR 200
90	3		AMPLOP FIS, FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA 200/36
89	3		FILM WARNA NEW SUPERIA 200/20, FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA 200/36
87	2		FILM WARNA HORI COLOR GX 100 135/36, FILM WARNA NEW SUPERIA X-TRA 400/36
86	2		FILM WARNA HORI COLOR GX 100 135/36, FILM FUJICOLOR 200
83	3		FILM WARNA HORI COLOR GX 100 135/36, FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA 200/36
81	2		CRYSTAL 80 5 X 610 G, AMPLOP FIS
81	2		CRYSTAL 80 5 X 610 G, FILM WARNA NEW SUPERIA 200/36
81	3		FILM WARNA HORI COLOR GX 100 135/36, FILM WARNA NEW SUPERIA X-TRA 400/36, FILM WARNA NEW SUPERIA 200/36
81	2		850123 CN-16Q NQ1R 10LTx2, FILM WARNA NEW SUPERIA 200/36
80	2		850123 CN-16Q NQ1R 10LTx2, AMPLOP FIS
80	3		CRYSTAL 80 5 X 610 G, AMPLOP FIS, FILM WARNA NEW SUPERIA 200/36
79	3		AMPLOP FIS, FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA X-TRA 400/36
78	3		850123 CN-16Q NQ1R 10LTx2, AMPLOP FIS, FILM WARNA NEW SUPERIA 200/36
77	3		FILM WARNA NEW SUPERIA 200/20, FILM FUJICOLOR 200, FILM WARNA NEW SUPERIA X-TRA 400/36
77	2		850006 CN-16Q NQ3R 8LTx2, 850123 CN-16Q NQ1R 10LTx2
77	2		850006 CN-16Q NQ3R 8LTx2, FILM WARNA NEW SUPERIA 200/36
76	2		850006 CN-16Q NQ3R 8LTx2, AMPLOP FIS
75	2		850045 CN-16Q NQ2R 8LTx2, FILM WARNA NEW SUPERIA 200/36
75	3		850006 CN-16Q NQ3R 8LTx2, AMPLOP FIS, FILM WARNA NEW SUPERIA 200/36

Gambar 6. Itemset yang Dihasilkan

Probability	Importance	Rule
1.000	1.178	KAMERA AVEC, FILM WARNA NEW SUPERIA X-TRA 400/36 -> SLING BAG
0.986	1.166	850045 CN-16Q NQ2R 8LTx2, 850006 CN-16Q NQ3R 8LTx2 -> 850123 CN-16Q NQ1R 10LTx2
0.981	1.168	KAMERA AVEC -> SLING BAG
0.980	1.123	KAMERA AVEC, FILM WARNA NEW SUPERIA 200/36 -> SLING BAG
0.974	1.072	CP47 P1R, FILM WARNA NEW SUPERIA 200/20 -> CP47 P2R
0.972	1.133	850045 CN-16Q NQ2R 8LTx2, AMPLOP FIS -> 850123 CN-16Q NQ1R 10LTx2
0.963	1.390	850006 CN-16Q NQ3R 8LTx2 -> 850123 CN-16Q NQ1R 10LTx2
0.962	1.102	CRYSTAL 80 6 X 610 L, CRYSTAL 80 8 X 295 L -> CP49E REPL
0.961	1.269	850006 CN-16Q NQ3R 8LTx2, FILM WARNA NEW SUPERIA 200/36 -> 850123 CN-16Q NQ1R 10LTx2
0.961	1.269	850045 CN-16Q NQ2R 8LTx2 -> 850123 CN-16Q NQ1R 10LTx2
0.961	1.236	850006 CN-16Q NQ3R 8LTx2, AMPLOP FIS -> 850123 CN-16Q NQ1R 10LTx2
0.960	1.205	850045 CN-16Q NQ2R 8LTx2, FILM WARNA NEW SUPERIA 200/36 -> 850123 CN-16Q NQ1R 10LTx2
0.959	1.280	CP47 P1R, FILM WARNA NEW SUPERIA X-TRA 400/36 -> CP47 P2R
0.957	1.250	CP47 P2R, FILM FUJICOLOR 200 -> CP47 P1R
0.957	1.142	VID+BATT+NS 400/36+FB, AMPLOP FIS -> RIO+BATT+NS 400/36+FB
0.957	1.193	CP47 P1R, 850006 CN-16Q NQ3R 8LTx2 -> CP47 P2R
0.957	1.193	CP47 P1R, 850123 CN-16Q NQ1R 10LTx2 -> CP47 P2R
0.956	1.194	CP47 P2R, 850045 CN-16Q NQ2R 8LTx2 -> CP47 P1R
0.956	1.167	CP47 P1R, 850045 CN-16Q NQ2R 8LTx2 -> CP47 P2R
0.952	1.058	VID+BATT+NS 400/36+FB, FILM FUJICOLOR 200 -> RIO+BATT+NS 400/36+FB
0.952	1.122	CP47 P2R, CRYSTAL 80 5 X 610 G -> CP47 P1R
0.949	1.005	VID+BATT+NS 400/36+FB, FILM WARNA NEW SUPERIA 200/20 -> RIO+BATT+NS 400/36+FB
0.948	1.676	CP47 P1R, FILM WARNA NEW SUPERIA 200/36 -> CP47 P2R
0.948	1.773	CP47 P2R, FILM WARNA NEW SUPERIA 200/36 -> CP47 P1R
0.948	1.676	CP47 P1R, AMPLOP FIS -> CP47 P2R
0.948	1.676	CP47 P1R -> CP47 P2R
0.946	1.306	850045 CN-16Q NQ2R 8LTx2, 850123 CN-16Q NQ1R 10LTx2 -> 850006 CN-16Q NQ3R 8LTx2
0.944	1.007	CP47 P2R, CRYSTAL 80 8 X 295 L -> CP47 P1R
0.944	1.205	850045 CN-16Q NQ2R 8LTx2, AMPLOP FIS -> 850006 CN-16Q NQ3R 8LTx2
0.940	1.304	CP47 P2R, FILM WARNA NEW SUPERIA X-TRA 400/36 -> CP47 P1R
0.938	1.209	CP47 P1R, FILM FUJICOLOR 200 -> CP47 P2R
0.937	1.083	850006 CN-16Q NQ3R 8LTx2, FILM WARNA NEW SUPERIA X-TRA 400/36 -> 850045 CN-16Q NQ2R 8LTx2
0.933	1.587	VID+BATT+NS 400/36+FB -> RIO+BATT+NS 400/36+FB

Gambar 7. Rules yang Dihasilkan

PENUTUP

Database yang menyimpan data transaksi penjualan tidak dapat mendukung analisa data dan pengambilan keputusan. Data transaksi penjualan yang ada dapat digunakan untuk analisis dengan menggunakan perangkat bantuan *business intelligence* yang ada dengan menentukan jumlah data yang dilibatkan dan penyesuaian terhadap format datanya. *Data warehouse* dapat digunakan untuk menyimpan data dalam jumlah besar dan juga berpotensi untuk mendapatkan sudut pandang baru dari sebaran data, dan memungkinkan untuk memberikan pelaporan dan jawaban dari pertanyaan pengguna yang bersifat ad hoc termasuk juga untuk melakukan analisis penjualan dengan konsep *cross-selling* untuk mendukung pengambilan keputusan.

Proses transformasi dari *database* yang menyimpan data transaksi penjualan mampu memindahkan data penting yang berguna bagi perusahaan untuk melakukan analisis. *Star Schema* yang telah dirancang dapat memenuhi fungsinya sebagai suatu tempat penampungan data yang memungkinkan para pengambil keputusan untuk pergi hanya ke satu tempat untuk mengakses seluruh data yang ada tentang organisasinya. Metode *cross-selling* dilakukan dengan memilih produk didasarkan pada frekuensi (seberapa sering) kemunculan produk selama periode transaksi berlangsung. Terdapat tiga indikator untuk menghasilkan kemungkinan *cross-selling* yang paling baik. *Support score*, *confidence score*, dan *improvement score*. Implementasi *data warehouse*, teknik *data mining* dan konsep *cross-selling* dapat membantu perusahaan menyimpan seluruh data data transaksi, melakukan analisis dan menghasilkan laporan yang berguna bagi perusahaan sebagai bahan pertimbangan untuk menganalisis suatu masaah dan mengambil keputusan.

DAFTAR PUSTAKA

- Buja, A., & Lee, Y. (2001). *Data mining criteria for tree-based regression and classification*, Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining San Francisco, CA, pp. 27-36.
- Chang, C. C., & Chen, R. S. (2006). Using Data Mining Technology to Solve Classification Problems: a Case Study of Campus Digital Library. *The Electronic Library* Vol. 24 No. 3, pp. 307-321, Emerald Group Publishing Limited.
- Chen, S. Y., & Liu, X. (2005). Data Mining from 1994 to 2004: an Application-Oriented Review. *International Journal of Business Intelligence and Data Mining*, Vol. 1 No. 1, pp. 4-11.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD Process for Extracting Useful Knowledge from Volumes Of Data. *Communications of the ACM*, Vol. 39 No. 11, pp. 7-34.
- Goodman, J. (1992). Leveraging the customer database to your competitive advantage. *Direct Marketing*, 55(8), 26– 27.
- Han, J., & Kamber, M. (2001). *Data Mining: Concepts and Techniques*. San Mateo, CA: Morgan Kaufmann.
- Hand, D. J. (1998). Data mining: statistics and more? *The American Statistician*, Vol. 52 No. 2, pp. 112-8.

- Kamakura, W. A., Wedel, M., Rosa, F., & Mazzon, J. A. (2003). Cross-Selling through Database Marketing: a Mixed Data Factor Analyzer for Data Augmentation and Prediction. *International Journal of Research in Marketing*, 20, 45-65.
- Mallach, E. G. (2000). *Decision Support and Data Warehouse Systems*. United States: McGraw-Hill.
- Wah, T. Y., Peng, N. H., & Hok, C. S. (2007). *Building Data Warehouse*. Proceedings of the 24th South East Asia Regional Computer Conference 2007, Bangkok, Thailand.
- Yu, S. C., & Chen, R. S. (2001). Developing an XML Framework for an Electronic Document Delivery System. *The Electronic Library*, Vol. 19 No. 2, pp. 102-10.