

INFRASTRUKTUR *BUSINESS INTELLIGENCE* MENDUKUNG *DATA MINING* DALAM PROSES *E-BUSINESS*

(Infrastructure of Business Intelligence to Support Data Mining for Company E-Business)

Violitta Yesmaya

Program Studi Teknologi Informasi
Universitas Bina Nusantara – Jakarta
vyesmaya@yahoo.com

Abstrak

Infrastruktur *business intelligence* (BI) merupakan suatu aplikasi berbasis teknik yang dapat membantu pemanfaatan data yang ada untuk menghasilkan suatu informasi baru yang bernilai. Dalam BI juga diperlukan kerangka yang terintegrasi dengan *knowledge* agar data-data yang dikumpulkan mengenai pelanggan memudahkan bagian manajerial untuk memberikan tanggapan yang tepat. *Business Intelligence-driven Data Mining* (BI_{DDM}), merupakan metode gabungan antara BI dengan *data mining*, yang menggabungkan dukungan *knowledge* dengan *data mining* dan metode yang didukung oleh *data mining*. Dalam membuat BI_{DDM} yang terdiri dari *Four-layer framework* ini diterapkan pada *business Intelligence* yang diperoleh dari metode *data mining*. BI_{DDM} dapat mendukung informasi *e-business* sehingga proses transaksi lebih mudah, akurat, dan dapat membantu pihak *top* manajerial dalam pengambilan keputusan secara tepat dan cepat.

Kata Kunci: *business intelligence, BI_{DDM}, data mining, e-business*

Abstract

Infrastructure of Business intelligence (BI) is a technology-based application useful for generating valuable new information based on the data available. BI also requires a framework integrated with the knowledge so that the data gathered about customers can help the managerial sections make appropriate responses. Business Intelligence-driven Data Mining (BI_{DDM}) is an integrated BI and data mining method which combines support for data mining and knowledge. The development BI_{DDM} includes Four-layer framework applied to Business Intelligence derived from the data mining method. BI_{DDM} potentially support accurate business information and help the top management to make accurate and quick decisions.

Keywords: *business intelligence, BI_{DDM}, data mining, four-layer framework*

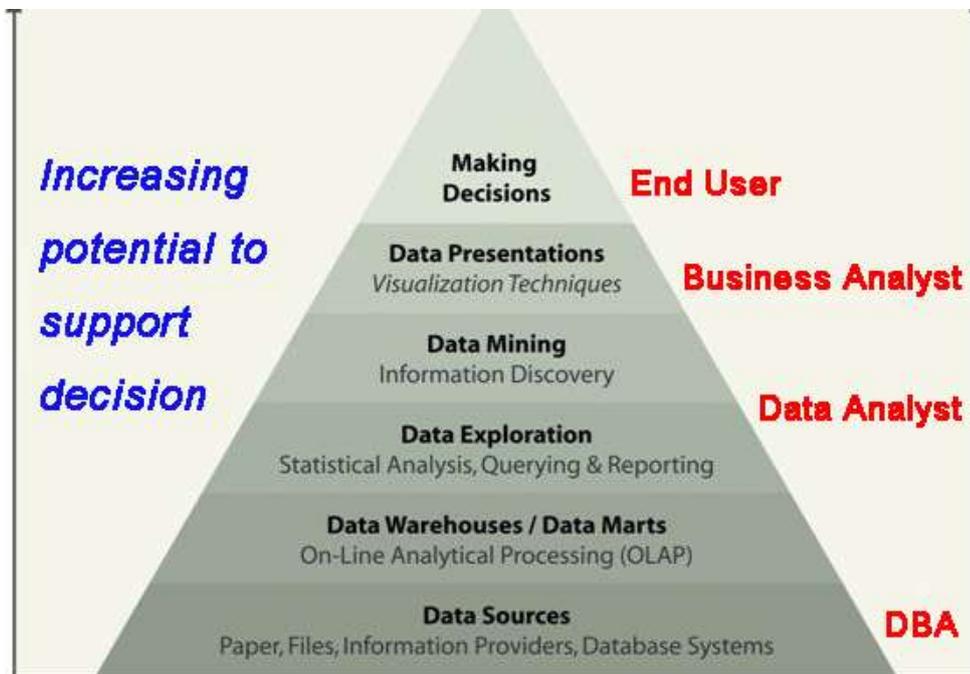
Tanggal Terima Naskah : 22 Februari 2013
Tanggal Persetujuan Naskah : 17 April 2013

1. PENDAHULUAN

Pada saat ini kebutuhan informasi yang cepat dan akurat dalam bidang bisnis merupakan hal yang penting dalam suatu perusahaan. Hal ini merupakan salah satu faktor yang mendukung bisnis di suatu perusahaan, agar menjadi lebih unggul dari perusahaan lain. Keunggulan yang dapat diperoleh adalah dengan mengetahui pengambilan keputusan yang tepat dan akurat oleh pihak *top* manajerial perusahaan. Dalam membuat keputusan, pihak *top* manajerial membutuhkan informasi yang jelas dan mudah dipahami,

sesuai dengan kebutuhan. Dalam mendukung informasi bisnis yang akurat maka mulai dikembangkan *e-business* (suatu kegiatan di internet, dengan menggunakan kolaborasi dan intrabisnis lainnya, lebih dari kegiatan jual beli pada umumnya), yang berkembang dari proses *e-service* [1]. *E-service* adalah suatu jaringan sistem komputer yang kompleks dan memiliki sejumlah data besar yang disimpan dalam *database*.

Business Intelligence (BI) merupakan suatu aplikasi *e-business* yang berguna untuk mengelola data-data dalam perusahaan (data operasional, transaksional, dan lain-lain) ke dalam suatu bentuk pengetahuan (informasi) yang dibutuhkan [2]. Dalam BI, sejumlah data besar yang berasal dari format dan sumber yang berbeda-beda dapat digabungkan dan dapat diubah untuk menjadi kunci pengetahuan bisnis. Pada Gambar 1 ditunjukkan gambaran umum mengenai cara suatu data ditransformasikan menjadi BI (*Business Intelligence*). Dalam proses transformasi BI melibatkan ahli bisnis dan juga beberapa ahli teknis, untuk dapat mengkonversikan data dalam jumlah yang besar. Ahli ini diperlukan agar menghasilkan hasil yang berarti sehingga dapat mendukung pengambilan keputusan yang dapat digunakan oleh *top* manajerial pada perusahaan.



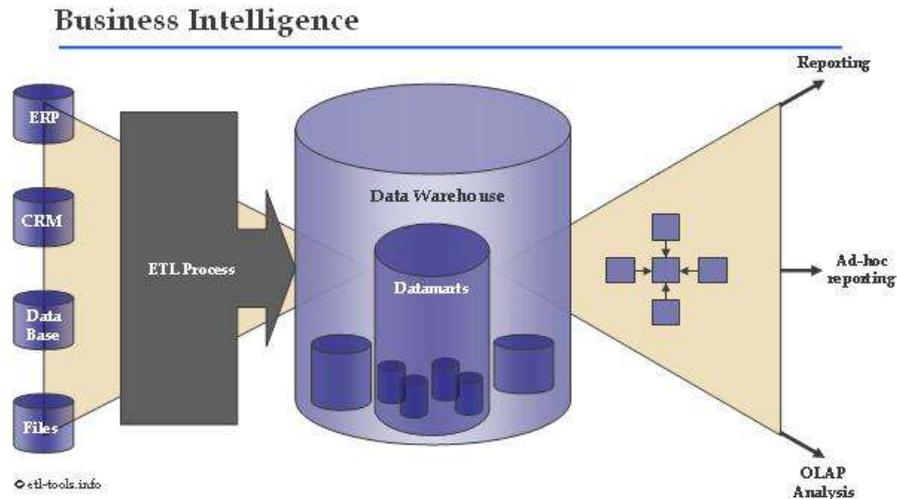
Gambar 1. *Business intelligence processing*

Pada proses BI terdapat beberapa tahapan, yaitu [2]:

- Bagian 1-2 (*data sources, data warehouse/data marts*): menjelaskan konsep yang berhubungan dengan BI dan DDM (*distributed data mining*) dimana permasalahan DDM dalam *e-business* dijelaskan pada bagian ini.
- Bagian 3 (*data exploration*): sistem DDM ditampilkan dalam tiga tampilan (*statistical analysis, querying, dan reporting*).
- Bagian 4 (*data mining*): mengusulkan sistem DDM yang sesuai dengan sumber data yang diperoleh, yang dapat mendukung hirarki dari data *mining*.
- Bagian 5 (*data presentations*): mengevaluasi kelayakan dari model DDM yang telah diusulkan pada bagian 4.
- Bagian 6 (*making decisions*): menampilkan kesimpulan dan saran hasil dari evaluasi, yang dapat berguna bagi penelitian di masa mendatang.

2. KONSEP DASAR

Business Intelligence (BI) merupakan aplikasi yang digunakan untuk mengumpulkan, menyimpan, menyediakan akses data. BI juga berfungsi untuk menganalisis data bisnis, seperti pendapatan dari hasil penjualan produk atau departemen, atau biaya yang berhubungan dengan pendapatan, yang dapat membantu pengguna bisnis dalam membuat pengambilan keputusan dengan lebih baik. BI umumnya memiliki kemampuan untuk menyediakan catatan di masa lalu, saat ini, dan prediksi dari operasi bisnis yang ada. Fungsi dari teknologi BI adalah pelaporan, *Online Analytical Processing*, *data mining*, kinerja manajemen bisnis, dan analisis prediksi [3].



Gambar 2. Proses BI

2.1 *Data Mining*

Data mining saat ini banyak digunakan di perusahaan-perusahaan. *Data mining* berfungsi agar terfokus pada konsumen, khususnya ritel, *finance*, komunikasi, dan organisasi pemasaran yang ada. Pemfokusan ini memungkinkan perusahaan untuk menentukan hubungan antar faktor-faktor "*internal*", yaitu harga, tata letak barang, dan faktor-faktor "*external*", seperti indikator ekonomi, persaingan, dan demografi pelanggan. Hal ini memungkinkan perusahaan untuk menentukan dampak penjualan, kepuasan pelanggan, dan keuntungan dari perusahaan. Pada akhirnya, perusahaan dapat menelusuri rangkuman informasi untuk melihat setiap detail transaksional. Dengan adanya *data mining*, ritel dapat dilihat dengan menggunakan catatan nilai jual dari setiap pembelian pelanggan sehingga dapat mengirimkan promosi-promosi yang ditargetkan berdasarkan catatan pembelian masing-masing pelanggan. Dengan *data mining* demografi yang didapat dari komentar atau kartu garansi, ritel dapat mengembangkan produk dan promosi untuk dapat menarik pelanggan-pelanggan tertentu.

Saat ini seiring dengan semakin berkembangnya informasi dimana transaksi dan sistem analisis sudah mulai terpisah, *data mining* menyediakan suatu hubungan antar keduanya (transaksi dan sistem analisis). Dalam *data mining* terdapat beberapa *software* analitis yang tersedia, yaitu statistik, *machine learning*, dan *neural networks* yang berguna untuk menyimpan data transaksi dari pelanggan. Terdapat beberapa teknik yang umum digunakan dalam *data mining*, yaitu [1]:

- *Classification*

Data yang tersimpan akan digunakan untuk mencari data dalam suatu kelompok yang telah ditentukan sebelumnya. Sebagai contoh, jaringan *bookstore* dapat menjadi sumber informasi data pembelian pelanggan untuk menentukan kapan pelanggan

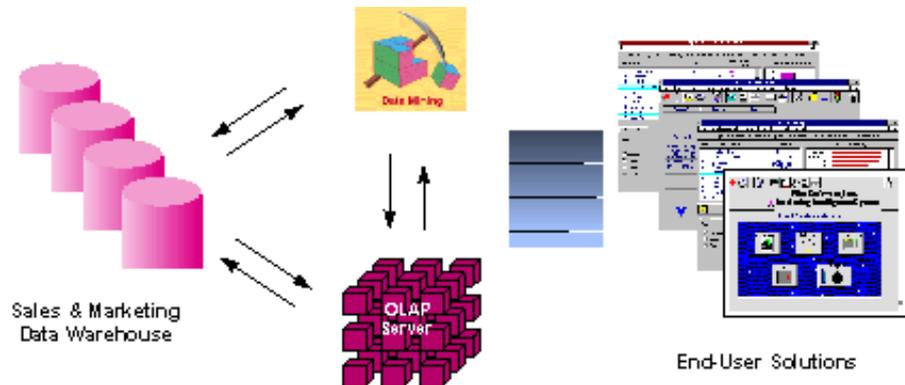
sering mengunjungi *bookstore* dan buku jenis apa yang biasanya dibeli di *bookstore*. Informasi ini dapat digunakan untuk meningkatkan perdagangan dengan memberikan promosi pada waktu-waktu tertentu.

- *Clustering*
Data item yang diidentifikasi berdasarkan hubungan yang logis atau yang sesuai dengan karakteristik tertentu. Algoritma *clustering* dapat digunakan untuk mengidentifikasi kelas pelanggan berdasarkan kebutuhan tertentu dari pelanggan tersebut. Sebagai contoh, data dapat digali dengan lebih dalam untuk mengidentifikasi segmen pasar mana yang sedang meningkat atau keterkaitan dengan konsumen.
- *Association*
Mengidentifikasi hubungan antara kejadian yang terjadi pada suatu waktu. Dalam ritel, terdapat upaya untuk mengidentifikasi produk apa yang mungkin laku di pasaran. Pada metode ini biasa dilakukan juga metode statistik. Misalnya, pelanggan membeli roti dan selai merupakan salah satu contoh *associative mining*.
- *Sequencing*
Teknik ini serupa dengan teknik *association*, yang berbeda adalah hubungan yang mungkin terjadi dalam suatu periode waktu (misalnya, kunjungan ke supermarket). Pembeli dapat dilacak dari pembelian yang terjadi karena pembeli dapat diidentifikasi dari nomor rekening atau cara yang lain. Sebagai contoh, ritel peralatan berkemah dapat memprediksi kemungkinan ransel yang dibeli berdasarkan pada pembelian *sleeping bag* dan sepatu *hiking*.
- *Regression*
Data dapat menjadi suatu sumber informasi untuk memprediksi nilai di masa mendatang atau kejadian yang mungkin terjadi di masa depan berdasarkan pada sumber informasi yang terdahulu atau hasil dari data statik.
- *Forecasting*
Estimasi nilai pada masa depan berdasarkan pola dari hasil data yang telah dikumpulkan (contohnya, peramalan permintaan suatu barang). Hal ini merupakan upaya untuk memanfaatkan waktu di masa depan dengan memprediksi penjualan di masa mendatang.
- Teknik lain
Biasanya dengan menggunakan metode *artificial intelligence*, yang menggunakan penalaran, *fuzzy logic*, *genetic algorithms*, *fractal-based transform*.
Dalam *data mining* juga terdapat lima elemen utama, yaitu:
 - *Extract, transform, dan load* dari transaksi data ke sistem *data warehouse*.
 - Menyimpan dan mengelola data dalam sistem *database* multidimensi.
 - Menyediakan akses data untuk analisis bisnis dan teknologi informasi yang profesional.
 - Analisis data dengan *software* aplikasi.
 - Menyajikan data dalam format yang sesuai dan mudah dimengerti oleh penggunanya, seperti grafik atau tabel.

2.2 Arsitektur Data Mining

Saat ini banyak peralatan *data mining* yang telah beroperasi di luar *data warehouse*, dimana diperlukan langkah-langkah lebih dalam mengekstraksi, mengimpor, dan menganalisis data. Hasil analisis *data warehouse* dapat diterapkan untuk meningkatkan proses bisnis pada organisasi perusahaan tersebut, khususnya dalam bidang manajemen promosi dan peluncuran barang baru. Gambar 3 menjelaskan mengenai arsitektur dalam analisis lanjutan pada *data warehouse*. Titik awal yang baik berasal dari *data warehouse* yang berisi kombinasi data internal untuk melacak semua kontak

pelanggan yang berhubungan dengan data pasar mengenai aktivitas pesaing yang sedang terjadi. Latar belakang informasi pada pelanggan juga merupakan potensi yang kuat untuk menjadi seorang calon pelanggan. *Warehouse* dapat diimplementasikan dalam berbagai sistem *database* relasional, seperti *Sysbase*, *Oracle*, *Redbrick*, dan lainnya, yang nantinya harus dioptimalkan untuk akses data yang fleksibel dan cepat [4].



Gambar 3. *Integrated data mining architecture*

2.3 *Online Analytical Processing*

Online Analytical Processing (OLAP) merupakan kunci utama dalam pembuatan BI. OLAP digunakan untuk dapat meningkatkan analisis bisnis dimana berguna bagi perhitungan *Decision Support System* (DSS) yang dilakukan oleh *top* manajerial perusahaan. Struktur OLAP yang multidimensi juga memungkinkan pengguna untuk dapat menganalisis data yang ingin dilihat dari bisnis yang ada, merangkum dengan batasan-batasan yang telah ditetapkan sebelumnya, seperti batasan produk, wilayah, dan sudut pandang dari bisnis pengguna. *Data mining* harus terintegrasi dengan *data warehouse* dan OLAP untuk membuat ROI (*Return Of Investment*) agar berfokus pada analisis bisnis yang ada. Integrasi dengan *data warehouse* dapat menyebabkan keputusan operasi secara otomatis dilaksanakan dan diketahui apakah sesuai atau tidak. Dalam permasalahan bisnis *data mining* memiliki tujuan penting seperti mengetahui calon pelanggan dan pengoptimalan promosi.

2.4 *Business Intelligence*

Business Intelligence (BI) merupakan proses-proses dari perkembangan aplikasi *knowledge management* (KM) yang membantu mengubah data menjadi informasi yang berguna sehingga dapat menjadi suatu pemahaman untuk mendapatkan kegiatan bisnis yang efektif. BI yang didukung oleh *data mining* (BIDDM) merupakan suatu metode *data mining* baru yang menggabungkan pengetahuan yang didukung *data mining* dan metode yang didukung oleh *data mining*. Hal ini memiliki tujuan untuk mengisi celah antara pengetahuan *business intelligence* dalam *e-commerce* dan berbagai macam metode *data mining*. BIDDM berusaha untuk tetap fleksibel dalam pengimplementasian *data mining* dalam memenuhi kebutuhan bisnis.

BIDDM (*Business Intelligence-driven Data Mining*) merupakan suatu teknologi *data mining*, yang dapat menggabungkan dan menggunakan kembali komponen dari metodologi *data mining* untuk membantu dalam pembuatan BI. Hal ini bertujuan agar mempercepat proses *data mining*, karena dalam *e-service* memiliki beberapa tahapan yang tergabung antara *knowledge* dan model *data mining*. Suatu *knowledge* yang didukung oleh *data mining* dapat digabungkan menjadi suatu proses yang terintegrasi.

2.5 Data Mining Terdistribusi

Data mining terdistribusi (DDM – *Distributed Data Mining*) merupakan sekumpulan *data mining* yang terdistribusi [5], yang disimpan dalam *database* lokal yang berbeda-beda. Pada desain *data mining* terdistribusi (DDM) digunakan pendekatan *bottom-up*, yang berkaitan dengan fragmentasi data, replikasi, komunikasi, dan pengintegrasian tugas [6]. Dalam *data mining* terdistribusi terdapat beberapa tantangan yang mencakup penerapan *output* yang akan dicapai dan algoritma yang akan digunakan agar mendapatkan hasil yang berarti. Dalam *data mining*, jika data yang terdistribusi berada dalam format yang berbeda-beda, maka dapat membuat kerusakan pada data dan menghalangi sistem DDM (*Distributed Data Mining*) itu sendiri.

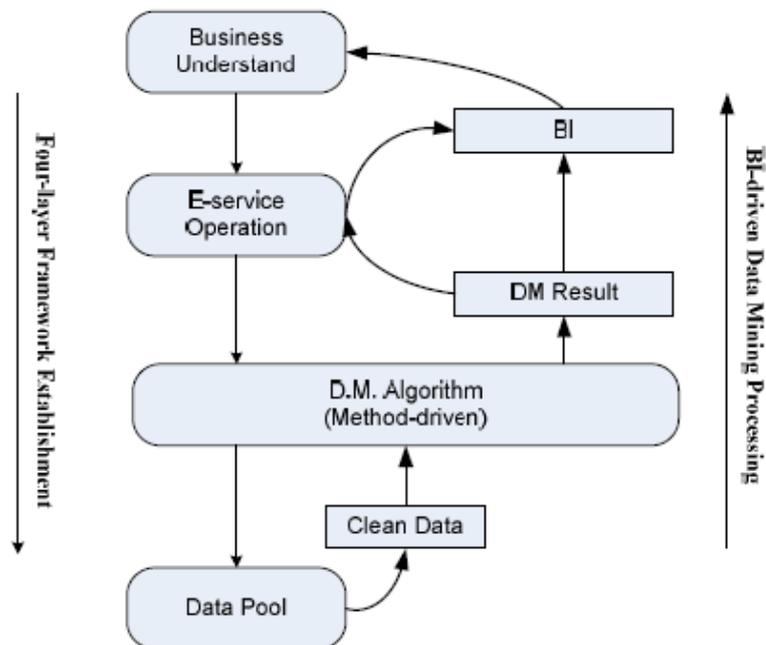
2.6 Data Mining yang Didorong Metode dan Pengetahuan

Data mining yang didukung oleh pengetahuan merupakan perhatian dalam *data mining*, diarahkan pada jenis pengetahuan yang dianggap berguna dibandingkan dengan menggunakan teknik atau algoritma [7]. Hal ini menyatakan bahwa kebutuhan *data mining* lebih baik menggunakan pengetahuan agar mencapai hasil yang lebih baik di masa mendatang.

Data mining yang didorong oleh metode diimplementasikan pada desain pendekatan *bottom-up*. *Data mining* yang didorong oleh pengetahuan mengandalkan pengetahuan dari para ahli dalam bidang yang sesuai, selama pemodelan dan *mining*.

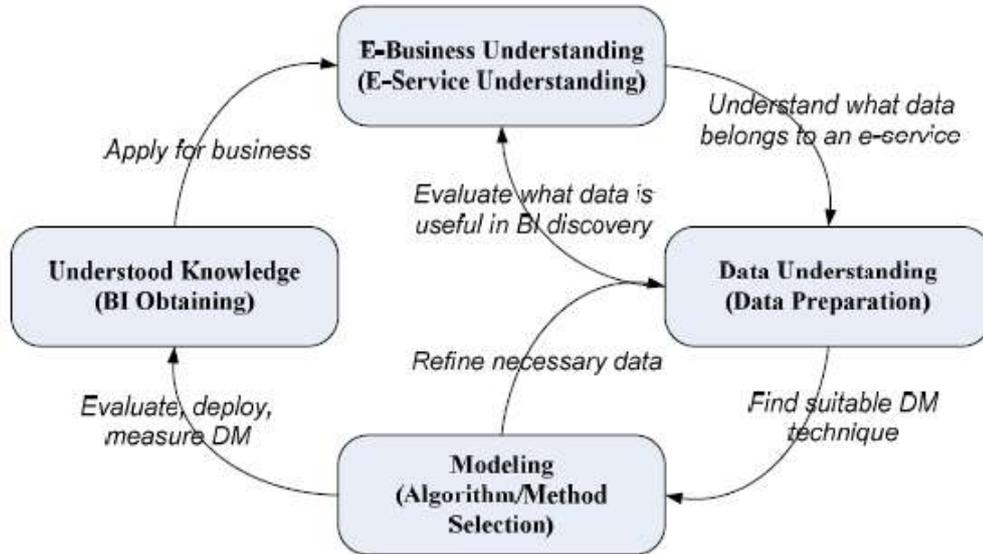
3. HASIL DAN PEMBAHASAN

Dalam BI_{DDM} (*Business Intelligence-driven Data Mining*) terdapat dua proses yang bekerja di dalamnya, yaitu konstruksi *framework* dan proses *data mining*. Konstruksi *framework* adalah proses untuk membentuk *four-layer framework*, dengan menggunakan pendekatan *top-down* dari BI ke dalam data. Pada proses *data mining* menggunakan pendekatan *bottom-up* untuk memudahkan penemuan *knowledge* yang berpotensi dari sumber data yang ada.



Gambar 4. Proses *flowchart* secara keseluruhan

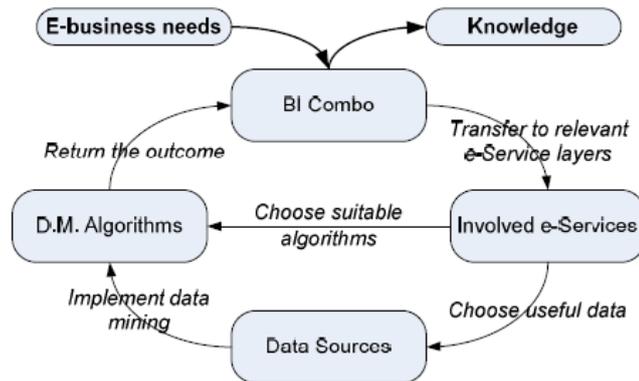
Proses konstruksi *framework* adalah proses dimana *four-layer* tahapan *framework* (dijelaskan pada Gambar 5) dibangun. Pada proses *top-down*, bagian *top* berguna untuk memahami bisnis dengan *e-service* yang berhubungan dengan bisnis dan sumber data yang berhubungan dengan *e-service*.



Gambar 5. Infrastruktur konstruksi *flowchart*

3.1 Proses *Data Mining* Dalam BI

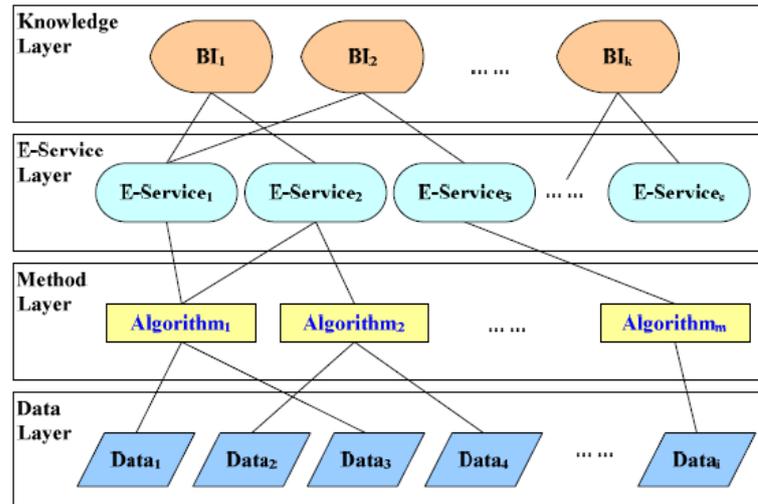
Ketika *four-layer framework* dibuat, maka satu atau beberapa algoritma pada *data mining* tertentu akan digunakan dalam BI bisnis pada perusahaan. Dalam pengertian ini, metode *data mining* dapat digunakan kembali karena dapat dihubungkan dengan BI lainnya. Hasil *data mining* disebut kombinasi BI, yang merupakan gabungan untuk menghasilkan pengetahuan bisnis yang baru. Proses ini ditunjukkan pada Gambar 6.



Gambar 6. *BI-driven data mining process flowchart*

3.2 *Four-Layer Framework*

BIDDM adalah suatu metodologi yang dibangun dari *four-layer framework* yang digambarkan pada Gambar 7. *Four-layer framework* ini diterapkan pada *business Intelligence* yang diperoleh dari metode *data mining*, tetapi BIDDM ini hanya sebagai proses *data query* dan *mapping*.



Gambar 7. Four-layer BI-driven data mining

Berikut ini adalah fungsi dari *four layer framework*:

- **Knowledge Layer**
 Layer pertama ini merupakan tahap awal dimana data atau informasi yang diperoleh sudah menjadi *knowledge* dan dapat mendukung pengambilan keputusan dalam bisnis. Dalam tahapan ini, pengguna akhir tidak perlu mengetahui mengenai rincian teknis pada penemuan BI dalam membuat keputusan bisnis yang strategis. Algoritma dan data yang dipilih akan dikendalikan oleh *e-service* dalam BIDDM. *Layer knowledge* ini mengikuti aplikasi *e-business*, seperti *e-payment*, *e-logistic*, *e-procurement*, *e-shopping*, *e-community*, dan segala sesuatu yang menggunakan satu atau lebih *e-service*. Dalam hal ini satu BI dapat berhubungan dengan lebih dari satu *e-service*.
- **Method Layer**
 Layer ini berisi mengenai algoritma *data mining*, yang digunakan untuk mengubah data ke dalam beberapa ekspresi yang berguna. Pada tahapan ini terdapat beberapa algoritma yang akan digunakan. Satu atau beberapa algoritma yang akan digunakan diharapkan nantinya akan menghasilkan olahan data yang akurat dan mudah dimengerti oleh pengguna.
- **Data Layer**
 Layer ini berada pada tahapan akhir yang bertanggung jawab untuk menyediakan sumber data dari *discovery knowledge*. Sumber data yang telah diproses sebelumnya akan digunakan untuk mengubah data mentah menjadi data matang. Data yang berguna datang dari sumber data mentah dengan menyaring data-data yang tidak konsisten. Dari hal ini dapat disimpulkan, *e-service* berguna dalam menghancurkan sistem data dari *data mining* yang tidak digunakan pada saat pendistribusian *data mining*. Dalam hal ini implementasi *data mining* terdistribusi lebih sulit dibandingkan dengan *data mining* terpusat karena permasalahan pada data yang sinkron atau data yang asinkron.

3.3 Knowledge Discovery Process

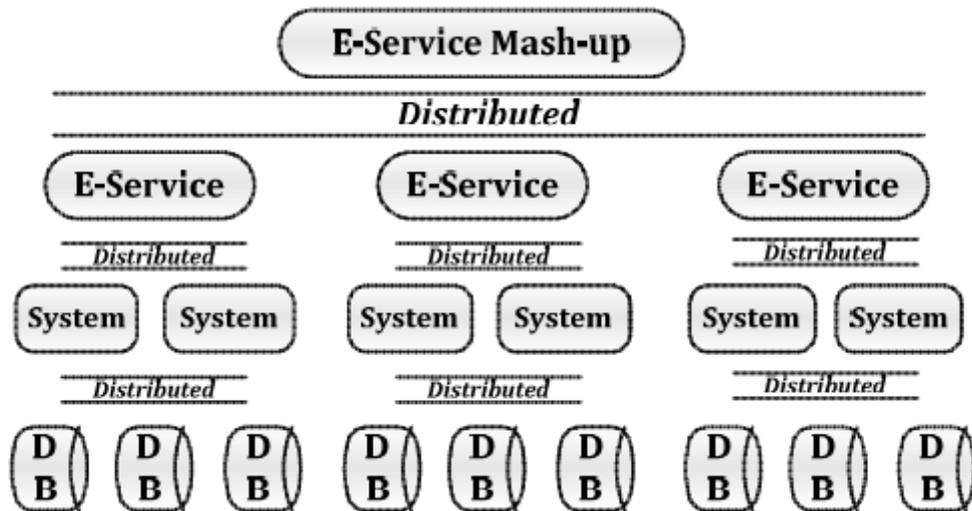
Knowledge discovery process (KDP) merupakan salah satu metode pencarian dalam aplikasi. KDP didefinisikan sebagai proses yang memiliki arti khusus dan mudah dimengerti oleh data [8]. KDP terdiri dari beberapa langkah, dimana dalam setiap

langkahnya berguna untuk menyelesaikan tugas-tugas tertentu. KDP ini juga berfokus dalam beberapa hal sebagai berikut:

- 1) Bagaimana data disimpan dan dapat diakses?
- 2) Bagaimana menggunakan algoritma yang efisien untuk menganalisis pengumpulan data yang besar?
- 3) Bagaimana menjelaskan dan menggambarkan hasil agar mudah dimengerti oleh pengguna?
- 4) Bagaimana model dan dukungan interaksi antara manusia dan mesin?

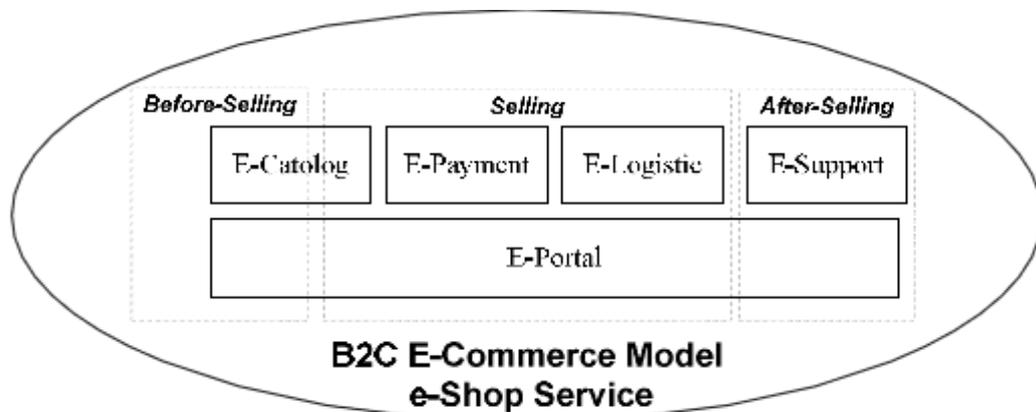
3.4 Arsitektur *E-Service* Terdistribusi

E-service berguna dalam mendistribusikan dan mentransfer informasi yang didapat melalui internet [9]. Pada setiap *e-service* terdapat tahapan awal, baik secara terdistribusi maupun secara terpusat, dimana di dalamnya terdapat *database* lokal yang menyediakan penyimpanan untuk sistem informasi yang sesuai dan merupakan sumber data bagi KDP.



Gambar 8. Arsitektur *E-service* terdistribusi

Misalnya terdapat sebuah model bisnis yang digunakan untuk menggambarkan B2C dan B2C (*Business to Customer*) pada *e-shop*. *E-shop* ini memiliki lima *e-service* dalam siklus bisnis secara keseluruhan seperti pada gambar 9.



Gambar 9. Contoh *e-service* dari B2C *e-shop*

Berikut ini adalah penjelasan masing-masing bagian:

- *e-portal* : sebuah *website* transaksi jual beli yang dapat menghasilkan data *log web* pada *web server*.
- *e-catalog* : daftar produk yang ada dalam *website*.
- *e-payment* : merupakan transaksi *online* untuk melakukan pembayaran. Proses ini biasanya dilakukan dengan tiga entitas, pelanggan, *e-shop*, dan pembayaran melalui *gateway* (*transfer, paypal*, dan lain-lain).
- *e-logistic* : proses untuk pengiriman barang ke pelanggan
- *e-support* : layanan yang disediakan untuk mendukung pencatatan penjualan, keluhan, servis, dan lain-lain.

4. KESIMPULAN

Dalam perkembangan *business intelligence*, BIdDM (*Business Intelligence-driven Data Mining*) merupakan suatu usulan yang sangat baik untuk *e-business* agar mendapatkan BI dengan data atau informasi yang jelas dan akurat, sehingga memudahkan *top* manajerial dalam pengambilan keputusan. Kegiatan *Business Intelligence* dapat memberikan nilai lebih bagi penggunanya, dimana proses BI ini dapat memudahkan perusahaan dalam menganalisis data dan transaksi para pelanggan.

BI yang didukung oleh *data mining* yang menggunakan *four-layer framework* membantu mengidentifikasi algoritma *mining* yang tepat agar menghasilkan *e-service* yang sesuai atau relevan. *Four-layer framework* merupakan tahapan-tahapan yang ada dalam BIdDM yang memiliki keunggulan untuk:

- Membuat keputusan bisnis menjadi strategis.
- Memiliki algoritma *data mining* yang dapat mengubah data menjadi suatu hasil olahan data yang akurat dan dapat dimengerti oleh penggunanya.
- Membantu mengubah sumber data mentah menjadi sumber data matang agar dapat menyaring data-data yang tidak konsisten.

REFERENSI

- [1]. Turban E, Aronson E. Jay, Ting-Peng Liang, "*Decision Support System and Intelligence System*," Seventh Edition, Prentice Hall, 2005.
- [2]. Bin Liu, Shi Gui Cao, Wu He, "*Distributed Data Mining for e-business*," Information Technology Management, 2001.
- [3]. Rainardi, Vincent, "*Building a Data Warehouse: With Examples in SQL Server*," Après, New York, 2008.
- [4]. Glenn J. Myatt, "*Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining*," John Wiley, ISBN: 0-470-07471-X, November 2006.
- [5]. Tibor S., Iveta Z., Lenka L., "*Distributed data mining and data warehouse*," XXX.ASR 2005 Seminar, Instruments and Control, 2005.
- [6]. Fu Y., "*Distributed Data Mining: An overview*," Newsletter of the IEEE Technical Committee on Distributed Processing, Spring, 2001.
- [7]. Graco, W., Semenova, T., and Dubossarsky, E., "*Toward knowledgedriven data mining*," In Proceedings of the 2007 international Workshop on Domain Driven Data Mining, DDDM '07 ACM, New York, 2007.
- [8]. K. J. Cios, W. Pedrycz, and R. M. Swiniarski, "*Data Mining: A Knowledge Discovery Approach*," Chapter 2, Springer Press, 2007.

- [9]. Zimmermann, H., “*OSI Reference Model--The ISO Model of Architecture for Open Systems Interconnection*,” IEEE Transactions on Communications, Volume 28, Issue 4, 1980.