

Andi Akram Yusuf

5113100173

Tugas Resume Paper

**Kesimpulan Dari Paper Analisis Penerapan *Datawarehouse* dan *Data Mining*
Klaim Asuransi Kendaraan Pada PT ABC**

- Database yang digunakan oleh PT ABC saat ini belum mendukung pengambilan data secara bersama-sama (banyak) secara cepat.
- Database saat ini belum mendukung proses pengambilan data untuk melihat *report* secara cepat.
- Penulis mengambil keputusan PT ABC sebaiknya membuat *datawarehouse* yang terpisah dari database operasional (sehari-hari) sehingga mempermudah proses pengambilan data dan tidak mengganggu database operasional.
- *Datamining* yang penulis sarankan adalah *datamining* dengan model *decision tree*.
- *Datawarehouse* ini juga akan mempermudah divisi finance untuk melihat dan memprediksi klaim yang akan datang.

ANALISIS PENERAPAN DATAWAREHOUSE DAN DATA MINING KLAIM ASURANSI KENDARAAN PADA PT ABC

Nama : Dewi Eka Mustikawati Nurdi, ST

Alamat : Kp pagadungan Rt 01/06 no 36 desa purwasari

Kecamatan purwasari kabupaten karawang

Email : dewi_eka29@yahoo.com

ABSTRAK

Kegiatan perencanaan, pengambilan keputusan dan evaluasi dapat dilakukan dengan lebih baik jika PT ABC memiliki informasi yang lengkap, tepat dan akurat serta dapat diakses dengan cepat. Informasi yang dibutuhkan dapat di ekstrak dari data operasional yang tersimpan dalam database yang terintegrasi. Penelitian ini mengkaji ekstraksi data operasional ke dalam sebuah *datawarehouse* untuk kemudian dilanjutkan dengan kegiatan analisis data menggunakan teknik *data mining*. Hasil dari penelitian ini adalah sebuah *datawarehouse*, selanjutnya dilakukan analisa data menggunakan teknik *data mining* terhadap *datawarehouse* yang sudah dibangun. Hasil dari penerapan data mining adalah dihasilkannya model data mining yang tepat untuk PT ABC sehingga memudahkan manajemen untuk memprediksi apakah klaim pelanggan *diaccept*, *reject* atau masih *outstanding* dan membantu manajemen untuk memprediksi budget klaim yang harus dikeluarkan.

Kata kunci : *Data warehouse*, *Data Mining*

PENDAHULUAN

Latar Belakang Masalah

Sebuah perusahaan keuangan non bank yang bergerak dibidang jasa asuransi kerugian nasional, yang selanjutnya dalam penulisan ini disebut PT ABC untuk menjaga kerahasiaan identitas perusahaan. PT ABC menawarkan jasa penanggungan resiko terhadap tertanggungnya dan akan memberikan jaminan atas kerugian yang diderita oleh tertanggung, jaminan tersebut sesuai dengan klausula atau perjanjian antara PT ABC dengan tertanggung.

Managemen dan divisi – divisi yang terkait membutuhkan sistem yang bisa memberikan laporan untuk memberikan akses pada informasi secara lebih cepat, terinci dan menyediakan informasi yang spesifik karena *reporting syistem* yang ada pada saat ini hanya menggunakan aplikasi windows dan untuk memproses laporan tersebut membutuhkan waktu yang sangat lama sehingga dapat memperlambat database yang kemudian akan memperlambat produksi yang berjalan pada waktu yang bersamaan. Untuk mengoptimalkan data yang ada dalam perusahaan yang jumlahnya cukup besar, *data warehouse* sangat tepat digunakan dalam membantu perusahaan membuat sebuah laporan.

Outstanding akseptasi klaim tertanggung yang memakan waktu cukup lama menjadi masalah yang cukup besar pada PT ABC sehingga diperlukan *tools* yang dapat membantu divisi klaim dan manajemen mengatasi masalah tersebut. Selain itu, divisi *finance & accounting* membutuhkan suatu *tools* untuk menentukan seberapa besar budget yang harus dikeluarkan oleh perusahaan untuk masalah klaim tertanggung agar divisi *finance & accounting* dapat mempersiapkan pengaturan keuangan lebih optimal. *Data mining* sangat tepat digunakan untuk membantu PT ABC dalam membantu memprediksi akseptasi klaim tertanggung.

Identifikasi Masalah

- Bagaimana mengekstraksi data operasional kedalam *datawarehouse*?
- Bagaimana menganalisa model *data mining* yang akan digunakan di PT ABC dalam memprediksi klasifikasi data klaim asuransi?

Batasan Masalah

Data yang digunakan dalam penelitian ini adalah data periode bulan januari tahun 1999 sampai dengan bulan juni tahun 2009, sedangkan bisnis yang dibahas dalam penelitian ini adalah bisnis MBU (asuransi kendaraan).

Perumusan Masalah

- Bagaimana memperoleh reporting klaim yang akurat untuk PT ABC?
- Model *data mining* apa yang paling tepat digunakan untuk PT ABC?

Tujuan Penelitian

Tujuan dari penelitian ini adalah memperoleh model *data mining* yang tepat, yaitu yang mempunyai akurasi yang paling besar sehingga dapat digunakan untuk memprediksi klasifikasi data klaim PT ABC dengan tepat.

Manfaat Penelitian

- *Datawarehouse* memudahkan manajemen dalam melihat laporan perusahaan dan memudahkan marketing untuk melihat klaim cabang marketing tersebut.
- Menemukan model *data mining* yang tepat untuk PT ABC.
- *Data mining* mempermudah divisi klaim dalam memprediksi keputusan akseptasi klaim yang akan terjadi di masa mendatang.
- *Data mining* mempermudah manajemen dalam memprediksi nilai klaim yang akan terjadi di masa yang akan datang.

TINJAUAN PUSTAKA

Pengenalan Asuransi

Asuransi adalah perjanjian antara dua pihak (atau lebih) yaitu perusahaan asuransi (penanggung) dengan orang yang berasuransi (tertanggung). Isi perjanjiannya adalah:

- Tertanggung akan membayar premi
- Atas dasar pembayaran premi ini penanggung akan:
 - 1 Memberikan tertanggung penggantian jika akibat peristiwa yang tidak pasti di masa yang akan datang tertanggung mengalami kerugian financial
 - 2 Memberikan pembayaran berdasarkan meninggal atau hidupnya orang yang diasuransikan.

DataWarehouse

Datawarehouse diperkenalkan pada tahun 1988 oleh W.H.Inmon. *Datawarehouse* adalah penyimpanan fakta data secara teratur dan dapat diakses yang

berguna sebagai dasar pengambilan keputusan manajemen yang lebih baik. *Datawarehouse* memberikan akses pada informasi mengenai bisnis, produk dan pelanggan suatu perusahaan.

Data Mining

Apa yang dibutuhkan untuk melakukan *data mining*? *data mining* membutuhkan identifikasi masalah, bersamaan dengan pengumpulan data yang dapat membawa pada pemahaman yang lebih baik tentang pasar, dan model komputer untuk menyediakan analisis statistik atau yang lainnya.

Proses Standar Untuk *Data Mining*

- Pemahaman bisnis (*business understanding*)
- Pemahaman data (*data understanding*)
- Persiapan data (*data preparation*)
- Pembuatan model (*modeling*)
- Evaluasi (*evaluation*)
- Pelaksanaan (*deployment*) Teknik – Teknik *Data Mining*:

Berikut penjelasan teknik yang digunakan dalam *data mining* menggunakan *microsoft sql server 2005*:

- 1) Teknik *Naive Bayes*
- 2) Teknik *Decision Trees*
- 3) Teknik *Neural Network*

METODE PENELITIAN

Tempat dan Waktu Penelitian

Penelitian dilakukan langsung ke lapangan yaitu di sebuah instansi asuransi PT ABC. Data yang diambil adalah data klaim asuransi kendaraan dalam periode sepuluh tahun, yaitu tahun 1999 hingga juni 2009, sedangkan penelitian ini dilakukan sejak bulan januari 2009 hingga bulan juli 2009.

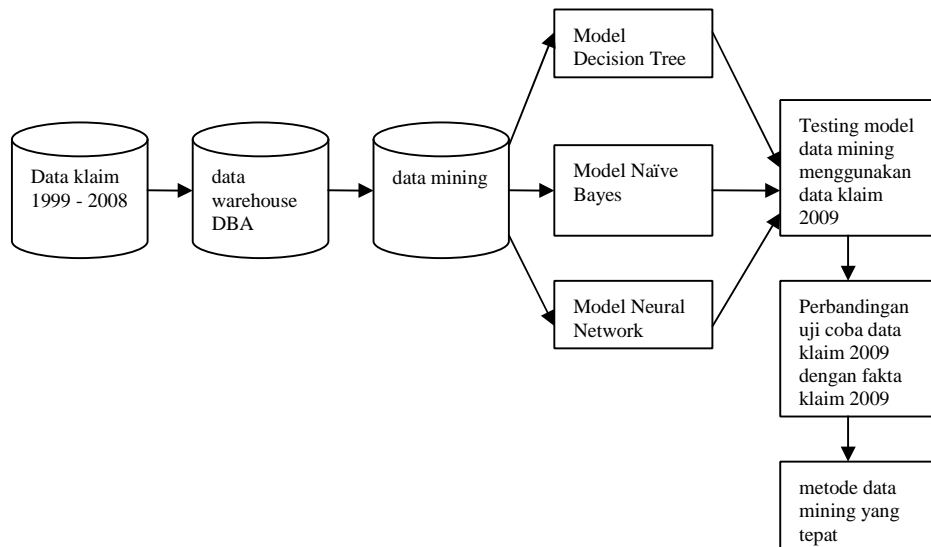
Metodologi Penelitian

- Studi pustaka

Penelitian ini menggunakan referensi seakurat mungkin untuk dapat memberikan hasil penulisan yang maksimal melalui telaah literatur seperti buku – buku, artikel serta jurnal.

- Studi Lapangan

1. teknik observasi, yaitu mengadakan pengamatan secara langsung terhadap objek yang akan diteliti. Pada gambar 3.1 dapat dijelaskan gambaran umum tahapan yang dilakukan dalam penelitian pada PT ABC.



Gambar 1. Tahapan analisis pemilihan model data mining

2. Teknik interview, yaitu mengadakan wawancara secara langsung dengan divisi IT yang terkait dengan data yang diambil.

Jenis Data Yang Diteliti

- Menurut cara memperolehnya: Data *primer* , yaitu data yang dikumpulkan langsung dari sumber pertama
- Menurut sumbernya: Data *internal*, yaitu data yang berasal dari dalam instansi mengenai kegiatan lembaga dan untuk kepentingan instansi itu sendiri

- Pembagian data menurut waktu pengumpulannya: Data *time series* adalah data yang dikumpulkan dari waktu - kewaktu pada satu obyek dengan tujuan untuk menggambarkan perkembangan
- Data menurut sifatnya: Data kualitatif adalah data yang berupa angka atau bilangan

Pengumpulan Data dan Informasi

Teknik tes digunakan untuk mengevaluasi data yang digunakan yaitu untuk mempertahankan keakuratan hasil laporan yang diperoleh untuk memperoleh hasil analisa yang akurat.

Analisis Sistem Yang Sedang Berjalan

Analisis sistem adalah proses pemecahan sistem menjadi beberapa sub sistem yang ruang lingkupnya lebih kecil, dengan maksud agar lebih mudah dalam mengidentifikasikan permasalahan – permasalahan, hambatan – hambatan, dan kesempatan – kesempatan yang ada dalam sistem, serta untuk mengetahui kebutuhan – kebutuhan sistem.

Analisis Permasalahan

Munculnya permasalahan mengindikasikan bahwa sistem yang ada belum memenuhi kebutuhan manajemen. Permasalahan tidak muncul begitu saja, melainkan ada faktor – faktor yang menyebabkan munculnya permasalahan tersebut. Untuk mengatasi suatu masalah, maka perlu diidentifikasi faktor- faktor yang menyebabkan munculnya permasalahan. Faktor- faktor ini bisa berasal dari dalam sistem itu sendiri (faktor *internal*) atau dari luar sistem (faktor *eksternal*).

Analisis Kebutuhan Sistem

Dari hasil analisis sistem dan analisis permasalahan pada sistem yang berjalan sekarang, maka sistem yang akan dibangun ini hendaknya mampu untuk menghasilkan output – output sebagai berikut:

1. Laporan klaim yang diterima, maupun ditolak sehingga pihak manajemen dapat mengetahui keuntungan atau kerugian perusahaan.
2. Memprediksi berapa jumlah klaim yang akan diterima atau ditolak atas klaim yang diajukan oleh pelanggan.

HASIL DAN PEMBAHASAN

Analisis Sistem

Sistem laporan produksi dan klaim yang digunakan pada PT ABC masih menggunakan aplikasi berbasis windows yaitu menggunakan aplikasi power builder 8.0, dimana tidak ada pembersihan yang dilakukan terhadap data – data yang di query sehingga proses *retrieve* data memakan waktu yang cukup lama karena data yang terdapat pada PT ABC mencapai jutaan data lebih.

Analisis Permasalahan

Aplikasi laporan produksi dan klaim berbasis windows pada saat ini tidak dilakukan proses pembersihan data terlebih dahulu sehingga membutuhkan waktu yang cukup lama jika user harus menarik data laporan, sedangkan laporan produksi dan klaim dibutuhkan tidak hanya untuk kantor pusat saja tapi setiap cabang PT ABC membutuhkan laporan tersebut. Sampai saat ini PT ABC memiliki 355 cabang di Indonesia. Jika proses dilakukan hampir bersamaan maka proses tersebut akan memperlambat database.

Selain faktor tersebut diatas, keterlambatan penanganan klaim akan berakibat buruk terhadap citra perusahaan sehingga perlu adanya suatu sistem yang membantu untuk penanggulangan masalah tersebut. Dengan adanya prediksi terhadap klaim kendaraan maka akan mudah dilakukan analisa klaim oleh departemen IT dan manajemen terhadap klaim tersebut.

Analisis Kebutuhan Prediksi

Beberapa faktor yang menyebabkan PT ABC membutuhkan prediksi akseptasi klaim yaitu pembayaran premi yang belum lunas, dokumen – dokumen klaim yang

belum lengkap, sistem yang kurang mendukung, sumber daya manusia, ketidakdisiplinan bengkel dalam memenuhi tanggung jawabnya.

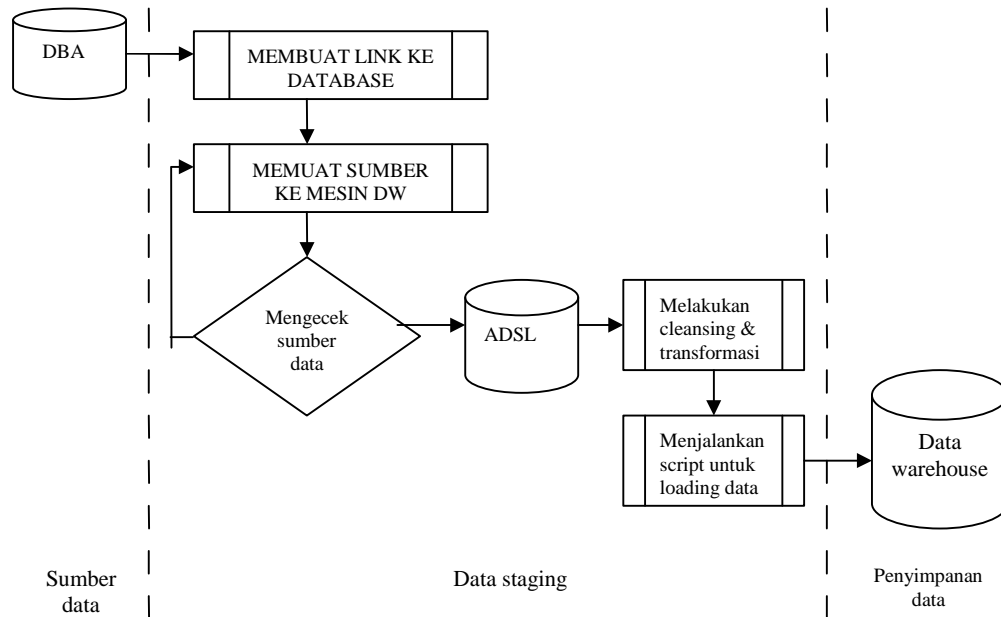
Selain faktor tersebut diatas, divisi *finance accounting* membutuhkan cek *budget* untuk memprediksi nilai yang akan dikeluarkan oleh *finance accounting*, semakin besar nilai klaim maka semakin besar pula budget yang harus dikeluarkan.

Perancangan Arsitektur *Data Warehouse*

1. Arsitektur Logikal

Rancangan arsitektur logikal dari *datawarehouse* PT ABC yang sekaligus menggambarkan proses pengisian data ke *datawarehouse* terdapat pada gambar 2, sumber data operasional yang digunakan adalah database DBA yang merupakan database yang berisi semua data PT ABC. Dari sumber data tersebut, secara periodik dilakukan pemilihan data dan selanjutnya dimuatkan ke database terpisah yaitu ADSL (disebut juga sebagai data *staging*) agar database operational tidak terganggu kinerja sistemnya.

Dalam proses pemuatan data ke database ADSL, selalu dilakukan proses pengecekan terlebih dahulu untuk memastikan bahwa data valid dan dapat dimuat ke database ADSL. Setelah itu, dilakukan proses pembersihan dan transformasi data pada database ADSL. Hasil dari proses pembersihan dan transformasi inilah yang kemudian disimpan ke dalam *datawarehouse*.



Gambar 2. Arsitektur logikal sistem *datawarehouse* dan *data mining*

2. Arsitektur Fisik

Rancangan arsitektur fisik dari *datawarehouse* PT ABC dapat dilihat pada gambar 3, pengguna mengakses *datawarehouse* melalui server aplikasi. Database ADSL dan *datawarehouse* berada pada mesin yang sama sehingga proses ekstraksi, transformasi dan loading dilakukan di mesin *datawarehouse* dan tidak mengganggu kinerja mesin operational.

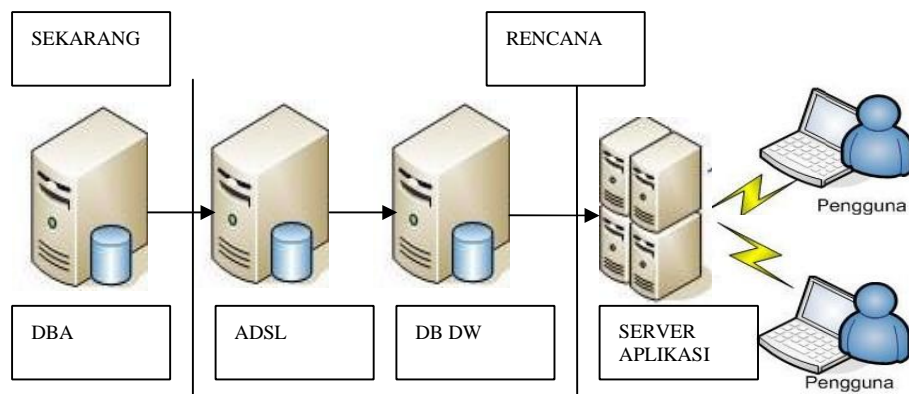
Basisdata yang digunakan adalah basisdata DBA yang merupakan *cluster* dari *engine* RX4000 dan RX6000, dimana masing-masing *engine* menggunakan *platform* Linux. Dari sumber data tersebut akan dilakukan proses pemilihan data dan transformasi yang kemudian akan di-*loading* ke dalam *datawarehouse* yang menggunakan *Microsoft SQL Server 2005*.

Data mining merupakan sebuah teknologi yang digunakan untuk mengambil keputusan. Dengan memanfaatkan *data mining* dengan bijaksana, keputusan yang diambil adalah *well-informed decision*. Selain tujuan diatas, *data mining* juga memiliki tujuan :

- Untuk mengolah data – data agregat
- Memakai sumber data yang berukuran sangat besar

- Memakai analisis

Alasan analisis inilah yang mendasari digunakannya *Microsoft SQL Server 2005* karena *Microsoft Corp* memakai frasa analisis servis untuk semua fasilitas – fasilitas di *Microsoft sql server 2005 DBMS*. Fasilitas analisis servis mencakup fasilitas *datawarehousing*, OLAP serta *data mining*.



Gambar 3. Arsitektur Fisik sistem *datawarehouse* dan *data mining*

Perencanaan Sumber Data

Sumber data yang digunakan berasal dari basisdata DBA untuk kurun waktu 1999-2009. Basisdata ini merupakan sistem basisdata terpusat yang terletak di kantor pusat. Basisdata ini berisikan data klien, data bengkel, data produksi, data klaim dan lainnya.

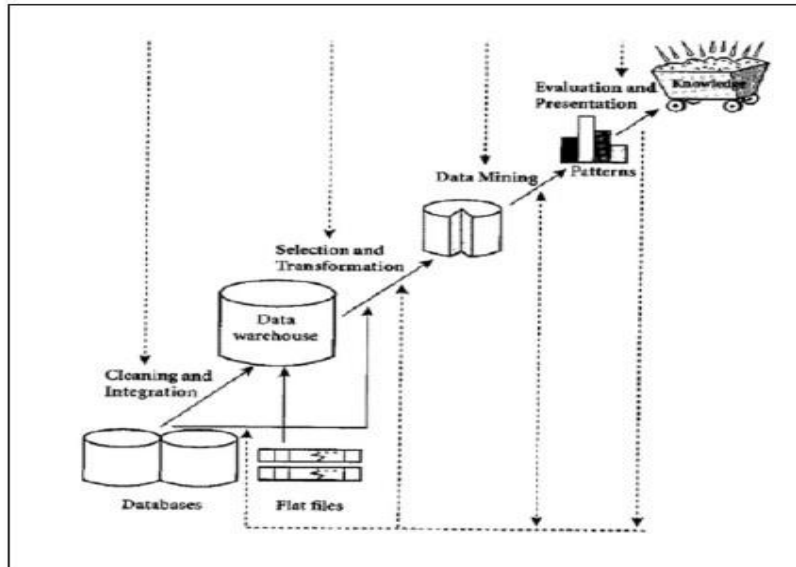
Basisdata ini memiliki total tabel 4904, total kolom 93254, jumlah baris 50 juta dengan kapasitas 500 GB. Sedangkan pertambahan data setiap harinya bisa mencapai 2 GB.

Pemodelan Data Dimensional

Skema yang digunakan untuk pemodelan data adalah *starflake schema* dimana terdapat satu tabel fakta dan beberapa tabel dimensi. Penggunaan *starflake schema* memungkinkan proses query yang lebih ringan dan memudahkan penjelajahan terhadap data dimensinya.

Analisis *Data mining*

1. Analisis Proses *data mining*



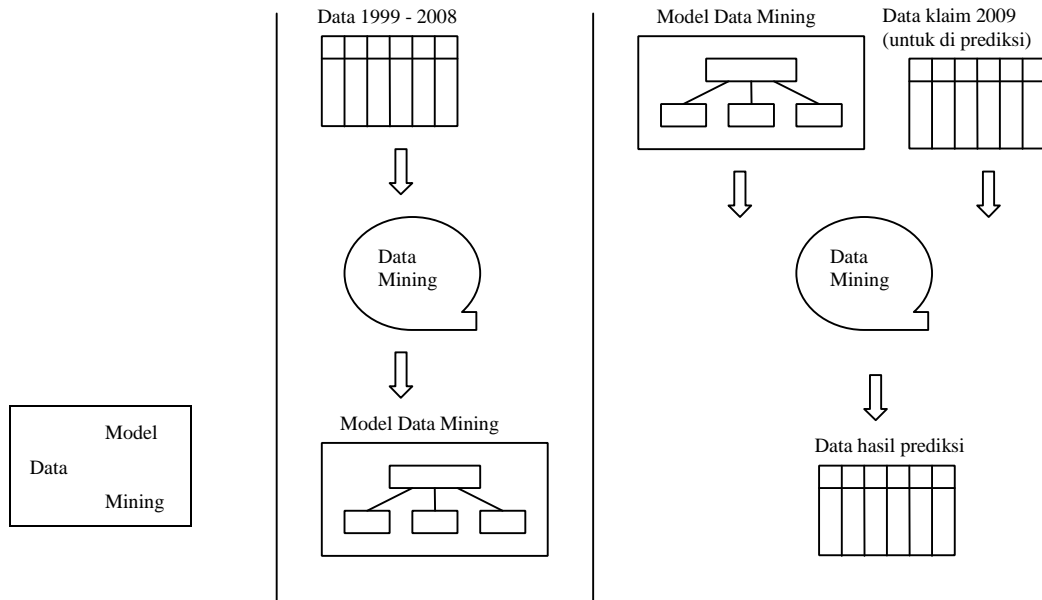
Gambar 4. Tahap – tahap *data mining*

Keterangan gambar 4:

Suatu rangkaian proses *data mining* dibagi menjadi beberapa tahap,yaitu :

- a. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*)
- b. Integrasi data (penggabungan data dari beberapa sumber)
- c. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-*mining*)
- d. Aplikasi teknik *data mining*
- e. Evaluasi pola yang ditemukan (untuk menemukan yang menarik / bernilai)
- f. Presentasi pengetahuan (dengan teknik visualiasi)

2. Analisis Model *Data Mining*



Gambar 5. Proses induksi, deduksi dan pengujian *data mining*

Keterangan gambar 5:

Dalam tahap membuat model *data mining*, terdapat tiga tahap yaitu:

- 1) Tahap pertama: model *data mining* mendefinisikan jumlah inputan kolom, memprediksikan kolom dan mengasosiasikan algoritma.
- 2) Tahap kedua: Algoritma *data mining* mulai menganalisa data yang diinput. Bergantung pada efisiensi sebuah algoritma, *data mining* memeriksa dataset kedalam datu atau lebih iterasi untuk menemukan korelasi antara nilai atribut. Pada tahap ini cukup memakan waktu. Setelah proses berhasil, model *data mining* dapat di browse menggunakan *content viewer* seperti *decision tree viewer*, *naive bayes viewer*.
- 3) Tahap ketiga: prediksi. Data klaim 2009 digunakan untuk mencoba memprediksi model *data mining*. Dalam tahap ini, *data mining* memasukkan aturan yang ditemukan dalam tahap kedua untuk dimasukkan dalam dataset dan menghasilkan prediksi.

3. Analisa Hasil Uji Coba *Data Mining*

Dalam tahap ini terdiri dari empat langkah yaitu:

- Pembuatan model *data mining*.
- Menguji coba model yang telah dibuat dengan memasukkan data klaim tahun 2009 yaitu dengan menggunakan mining *accuracy chart* dan *mining model prediction*, dimana tabel yang ingin diuji coba dimasukkan ke dalam model yang sudah dibuat pada tahap pertama lalu di filter berdasarkan prediksi yang ingin diketahui.
- Data yang sudah diuji tersebut di cek keakuratannya dengan memasukkan data klaim 2009 yang sudah memiliki status akseptasi.
- Kesimpulan dari hasil uji coba model *data mining*

KESIMPULAN DAN SARAN

Kesimpulan

Penelitian ini telah memperoleh hasil sebagai berikut:

- Dengan adanya *datawarehouse*, membuat proses penyusunan laporan cabang maupun kantor pusat terhadap klaim asuransi menjadi lebih mudah.
- Model yang paling tepat digunakan di PT ABC adalah model *decision tree*.
- Analisa data menggunakan *data mining*, mempermudah divisi klaim untuk memprediksi status akseptasi klaim.
- Analisa data menggunakan *data mining* mempermudah divisi *finance accounting* untuk memprediksi jumlah nilai klaim yang akan diterima.

Saran

- Mengembangkan model *data mining* yang belum di teliti seperti *Support Vector Machine (SVM)*, *Bayesian Network* dan lain – lain, sehingga diperoleh model *data mining* yang tepat untuk PT ABC karena perubahan bisnis bisa terjadi dalam setiap tahunnya.

DAFTAR PUSTAKA

- Bernaridho I Hutabarat. 2005. *Datawarehousing dengan SQL Server 2005*. Cetakan pertama. PT Elex Media Komputindi, Jakarta.
- David Olson, Yong shi. 2008. *Introduction to Business Data Mining*. Penerbit Salemba Empat, Jakarta.
- Gordon C A Dickson M Litt. 1984. *Introduction to insurance*. London
- Jozef Zurada and Subhash Lonia. 2005. *Comparison Of The Performance Of Several Data Mining Methods For Bad Debt Recovery In The Healthcare Industry*. Volume 21, Number 2.
- PT ABC. 2004. *Kebijakan Underwriting Motor Business Unit (MBU)*. No 002/M.ABC/II/2004.
- Reed jacobson, stacia Misner, Hitachi Consulting. 2005. *Microsoft SQL Server 2005 Analisis Services*. H.B. Fenn and Company Ltd, canada.
- Robert Sanderson. 2008. *Data Mining. Dept. of Computer Science University of Liverpool*.
- Vincent Rainardi. 2008. *Building a Data Warehouse: With Examples in SQL Server*. United States of America
- ZhaoHui Tang, Jamie MacLennan. 2005. *Data Mining with SQL Server 2005*. Wiley Publishing, Inc., Indianapolis, Indiana Published simultaneously in Canada
- Shamsul Chowdhury, Joseph O. Chan,cWalter E. Heller. 2007. *Warehousing and Data Mining: A Course in MBA and MSIS. College of Business Administration Roosevelt University*. Data Program from Uses Perspective2007 Volume 7 Issue 2