

APLIKASI DATA MINNING

Ari Fadli

fadli.te.unsoed@gmail

http://fadli84.wordpress.com

Lisensi Dokumen:

Copyright © 2003-2011 IlmuKomputer.Com

Seluruh dokumen di IlmuKomputer.Com dapat digunakan, dimodifikasi dan disebarkan secara bebas untuk tujuan bukan komersial (nonprofit), dengan syarat tidak menghapus atau merubah atribut penulis dan pernyataan copyright yang disertakan dalam setiap dokumen. Tidak diperbolehkan melakukan penulisan ulang, kecuali mendapatkan ijin terlebih dahulu dari IlmuKomputer.Com.

Pada tulisan kali ini penulis akan sedikit berbagi ilmu mengenai Konsep Dasar dari Data Mining atau biasa diterjemahkan kedalam bahasa Indonesia sebagai Menggali Informasi yang Terpendam

Pendahuluan

Sebagai cabang ilmu baru di bidang komputer (lihat artikel sebelumnya berjudul 'Data Mining') cukup banyak penerapan yang dapat dilakukann oleh Data Mining. Apalagi ditunjang ke-kaya-an dan ke-anekaragam-an berbagai bidang ilmu (artificial intelligence, database, statistik, pemodelan matematika, pengolahan citra dsb.) membuat penerapan data mining menjadi makin luas. Di bidang apa saja penerapan data mining dapat dilakukan? Artikel singkat ini berusaha memberikan jawabannya..

Analisa Pasar dan Manajemen

Untuk analisa pasar, banyak sekali sumber data yang dapat digunakan seperti transaksi kartu kredit, kartu anggota club tertentu, kupon diskon, keluhan pembeli, ditambah dengan studi tentang gaya hidup publik.

Beberapa solusi yang bisa diselesaikan dengan data mining diantaranya:

1. Menembak target pasar

Data mining dapat melakukan pengelompokan (clustering) dari model-model pembeli dan melakukan klasifikasi terhadap setiap pembeli sesuai dengan karakteristik yang diinginkan seperti kesukaan yang sama, tingkat penghasilan yang sama, kebiasaan membeli dan karakteristik lainnya.

2. Melihat pola beli pemakai dari waktu ke waktu

Data mining dapat digunakan untuk melihat pola beli seseorang dari waktu ke waktu. Sebagai contoh, ketika seseorang menikah bisa saja dia kemudian memutuskan pindah dari single account ke joint account (rekening bersama) dan kemudian setelah itu pola belinya berbeda dengan ketika dia masih bujangan.

3. Cross-Market Analysis

Kita dapat memanfaatkan data mining untuk melihat hubungan antara penjualan satu produk dengan produk lainnya. Berikut ini saya sajikan beberapa contoh:

- Cari pola penjualan Coca Cola sedemikian rupa sehingga kita dapat mengetahui barang apa sajakah yang harus kita sediakan untuk meningkatkan penjualan Coca Cola?
- Cari pola penjualan IndoMie sedemikian rupa sehingga kita dapat mengetahui barang apa saja yang juga dibeli oleh pembeli IndoMie. Dengan demikian kita bisa mengetahui dampak jika kita tidak lagi menjual IndoMie.

4. Profil Customer

Data mining dapat membantu Anda untuk melihat profil customer/pembeli/nasabah sehingga kita dapat mengetahui kelompok customer tertentu suka membeli produk apa saja.

a. Identifikasi Kebutuhan Customer

Anda dapat mengidentifikasi produk-produk apa saja yang terbaik untuk tiap kelompok customer dan menyusun faktor-faktor apa saja yang kira-kira dapat menarik customer baru untuk bergabung/membeli.

b. Menilai Loyalitas Customer

VISA International Spanyol menggunakan data mining untuk melihat kesuksesan program-program customer loyalty mereka. Anda bisa lihat di www.visa.es/ingles/info/300300.html

c. Informasi Summary

Anda juga dapat memanfaatkan data mining untuk membuat laporan summary yang bersifat multi-dimensi dan dilengkapi dengan informasi statistik lainnya.

Analisa Perusahaan dan Manajemen Resiko

a. Perencanaan Keuangan dan Evaluasi Aset

Data Mining dapat membantu Anda untuk melakukan analisis dan prediksi cash flow serta melakukan contingent claim analysis untuk mengevaluasi aset. Selain itu Anda juga dapat menggunakannya untuk analisis trend.

b. Perencanaan Sumber Daya (Resource Planning)

Dengan melihat informasi ringkas (summary) serta pola pembelanjaan dan pemasukan dari masing-masing resource, Anda dapat memanfaatkannya untuk melakukan resource planning.

c. Persaingan (Competition)

o Sekarang ini banyak perusahaan yang berupaya untuk dapat melakukan competitive intelligence. Data Mining dapat membantu Anda untuk memonitor pesaing-pesaing Anda dan melihat market direction mereka.

o Anda juga dapat melakukan pengelompokan customer Anda dan memberikan variasi harga/layanan/bonus untuk masing-masing grup.

o Menyusun strategi penetapan harga di pasar yang sangat kompetitif. Hal ini diterapkan oleh perusahaan minyak REPSOL di Spanyol dalam menetapkan harga jual gas di pasaran.

d. Telekomunikasi

Sebuah perusahaan telekomunikasi menerapkan data mining untuk melihat dari jutaan transaksi yang masuk, transaksi mana sajakah yang masih harus ditangani secara manual (dilayani oleh orang). Tujuannya tidak lain adalah untuk menambah layanan otomatis khusus untuk transaksi-transaksi yang masih dilayani secara manual. Dengan demikian jumlah operator penerima transaksi manual tetap bisa ditekan minimal.

e. Keuangan

Financial Crimes Enforcement Network di Amerika Serikat baru-baru ini menggunakan data mining untuk me-nambang trilyunan dari berbagai subyek seperti property, rekening bank dan transaksi keuangan lainnya untuk mendeteksi transaksi-transaksi keuangan yang mencurigakan (seperti money laundry). Mereka menyatakan bahwa hal tersebut akan susah dilakukan jika menggunakan analisis standar. Anda bisa lihat di www.senate.gov/~appropriations/treasury/testimony/sloan.htm. Mungkin sudah saatnya juga Badan Pemeriksa Keuangan Republik Indonesia menggunakan teknologi ini untuk mendeteksi aliran dana BLBI.

f. Asuransi

Australian Health Insurance Commision menggunakan data mining untuk mengidentifikasi layanan kesehatan yang sebenarnya tidak perlu tetapi tetap dilakukan oleh peserta asuransi. Hasilnya? Mereka berhasil menghemat satu juta dollar per tahunnya. Anda bisa lihat di www.informationtimes.com.au/data-sum.htm. Tentu saja ini tidak hanya bisa diterapkan untuk asuransi kesehatan, tetapi juga untuk berbagai jenis asuransi lainnya.

g. Olah Raga

IBM Advanced Scout menggunakan data mining untuk menganalisis statistik permainan NBA (jumlah shots blocked, assists dan fouls) dalam rangka mencapai keunggulan bersaing (competitive advantage) untuk tim New York Knicks dan Miami Heat.

h. Astronomi

Jet Propulsion Laboratory (JPL) di Pasadena, California dan Palomar Observatory berhasil menemukan 22 quasar dengan bantuan data mining. Hal ini merupakan salah satu kesuksesan penerapan data mining di bidang astronomi dan ilmu ruang angkasa. Anda bisa lihat di www.aig.jpl.nasa.gov/public/mls/news/SKICAT-PR12-95.html.

i. Internet Web Surf-Aid

IBM Surf-Aid menggunakan algoritma data mining untuk mendata akses halaman Web khususnya yang berkaitan dengan pemasaran guna melihat perilaku dan minat customer serta melihat ke-efektif-an pemasaran melalui Web.

Dengan melihat beberapa aplikasi yang telah disebutkan di atas, terlihat sekali potensi besar dari penerapan Data Mining di berbagai bidang. Bahkan beberapa pihak berani menyatakan bahwa Data Mining merupakan salah satu aktifitas di bidang perangkat lunak yang dapat memberikan ROI (return on investment) yang tinggi. Namun demikian, perlu diingat bahwa Data Mining hanya melihat keteraturan atau pola dari sejarah, tetapi tetap saja sejarah tidak sama dengan masa datang. Contoh: jika orang terlalu banyak minum Coca Cola bukan berarti dia pasti akan kegemukan, jika orang terlalu banyak merokok bukan berarti dia pasti akan kena kanker paru-paru atau mati muda. Bagaimanapun juga data mining tetaplah hanya alat bantu yang dapat membantu manusia untuk melihat pola, menganalisis trend dsb. dalam rangka mempercepat pembuatan keputusan.

Data Mining dan Web Mining

Data mining (DM) yang juga dikenal sebagai Knowledge Discovery (Frawley et al., 1992) , merupakan salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database skala besar yang makin banyak terakumulasi sejalan dengan pertumbuhan teknologi informasi. Secara umum, data mining dapat didefinisikan sebagai suatu rangkaian proses untuk menggali nilai tambah berupa ilmu pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data (Pramudiono, 2003).

Web mining merupakan penerapan teknik data mining terhadap web dengan tujuan untuk memperoleh pengetahuan dan informasi lebih dari dalam web. Web mining dapat dikategorikan ke dalam tiga ruang lingkup yang berbeda, yaitu web content mining, web structure mining dan web usage mining (Srivastava et al., 2000).

Association Rules dan Algoritma Apriori

Association rules merupakan salah satu teknik data mining yang berfungsi untuk menemukan asosiasi antar variabel, korelasi atau suatu struktur diantara item atau objek-objek didalam database transaksi, database relasional, maupun pada penyimpanan informasi lainnya.

Sebagai ilustrasi dalam analisis weblog dari association rules adalah sebagai berikut, pola yang mungkin adalah “jika seseorang mengunjungi website CNN, terdapat kemungkinan sebesar 60% orang tersebut mengunjungi website Detik pada bulan yang sama.” Pada ilustrasi tersebut, pola yang ditemukan berpotensi menghasilkan potongan informasi yang menarik dan dibutuhkan oleh perusahaan yang terkait.

Proses di dalam teknik association rules adalah mencari aturan-aturan yang memenuhi minimum support dan confidence. Algoritma yang pertama kali digunakan dalam teknik association rules dan yang paling banyak digunakan adalah algoritma apriori (Agrawal & Srikant, 1994).

Web Crawler

Web crawler (yang juga dikenal dengan web spider atau web robot) adalah suatu program atau script otomatis yang menjelajahi WWW dengan menggunakan sebuah metode atau cara yang otomatis. Nama-nama yang jarang digunakan pada sebuah web crawler adalah ants, automatic indexers, bots, worms (Kobayashi & Takeda, 2000).

Extended Log File Format

Extended Log Format dirancang untuk memenuhi beberapa kebutuhan di bawah ini (Baker & Behlendorf, 1996):

- a. Memperbolehkan kontrol pada data yang direkam.
- b. Memenuhi kebutuhan proxy, client dan server dalam format yang umum.
- c. Menyediakan penanganan yang sempurna akan masalah penghilangan karakter.
- d. Memperbolehkan dalam pertukaran demografis data. Memperbolehkan dalam menyajikan rekapitulasi data.

Biografi Penulis



Ari Fadli, Lahir di Cirebon, 31 Juli 1984. Menamatkan SMU di SMU Negeri 4 Cirebon. Menyelesaikan program S1 dari Jurusan Teknik Program Studi Teknik Elektro, Universitas Jenderal Soedirman Puwokerto pada tahun 2007. Saat ini menjadi dosen di Jurusan Teknik Program Studi Teknik Elektro, Universitas Jenderal Soedirman Puwokerto dan sedang menyelesaikan studi di pascasarjana universitas Gadjah Mada Jurusan Teknik Elektro dengan Spesifikasi Sistem Komputer dan Informasi . Kompetensi awalnya adalah bidang basis data, Sistem informasi, sistem pakar dan saat ini sedang bergerak ke arah open source