

ANALISA ASOSIATIF DATA MINING UNTUK MENGETAHUI POLA KECELAKAAN LALU LINTAS

Agus Sasmito Aribowo

Prodi Teknik Informatika UPN "Veteran" Yogyakarta
Jl. Babarsari 2 Tambakbayan 55281 Telp (0274) 485323
email : sasmito_skom@yahoo.com

Abstract

The data of traffic accident can be processed into information that is important for Police Department. Those important information researched is to analyze the traffic accident data to find out is there any link between the occurrence of an accident to a certain brand of vehicle.

This research implementing data mining method to process the data traffic accident by using data mining techniques called Apriori Method. Apriori Method is used to identify a pattern of accidents based on brand, type of vehicles, and the vehicle's color. The results are used to estimate whether there is any correlation between the occurrences of a traffic accident to a particular brand.

The result can help the Police Department to find out whether there is any correlation between the occurrence of traffic accidents to the brand, type and the color of vehicle.

Keywords : rule based classification, apriori, brand loyalty, traffic accident

Data kecelakaan lalu lintas dapat diolah menjadi informasi yang penting bagi Kepolisian. Informasi diperoleh dengan cara menganalisis data kecelakaan lalu lintas untuk mengetahui apakah ada hubungan antara terjadinya kecelakaan untuk merek kendaraan dan jenis kendaraan. Penelitian ini menerapkan metode data mining untuk memproses kecelakaan lalu lintas data dengan menggunakan teknik data mining yang disebut Metode Apriori. Apriori Metode digunakan untuk mengidentifikasi pola kecelakaan berdasarkan merek dan jenis kendaraan. Hasilnya digunakan untuk memperkirakan apakah ada korelasi antara kejadian kecelakaan lalu lintas dengan merek dan jenis kendaraan tertentu. Hasilnya dapat membantu Kepolisian untuk mengetahui apakah ada korelasi antara kejadian kecelakaan lalu lintas dengan jenis dan merek kendaraan.

Kata kunci : aturan klasifikasi dasar, apriori, brand loyalty, tingkat kecelakaan

1. PENDAHULUAN

Salah satu bentuk interaksi yang bersifat kausal dan saling membutuhkan adalah interaksi antara dealer penjualan kendaraan bermotor dengan Kepolisian Republik Indonesia (POLRI). Dealer kendaraan bermotor berperan menyumbangkan data kendaraan bermotor kepada pihak Kepolisian. Kepolisian mencatat penggunaan kendaraan tersebut, diantaranya adalah pendataan kecelakaan. Pendataan

Salah satu bentuk informasi penting yang dibutuhkan oleh kepolisian adalah pola kecelakaan lalu lintas. Data kecelakaan lalu lintas menggambarkan kelalaian dalam penggunaan kendaraan bermotor. Data tersebut juga dapat menggambarkan kenyamanan dan kelayakan kondisi suatu kendaraan. Analisa dibutuhkan untuk mengetahui mengapa suatu kecelakaan dapat terjadi dan dapat diarahkan untuk meneliti adakah keterkaitan antara terjadinya suatu kecelakaan dengan merek, tipe dan warna kendaraan.

Penelitian ini berlokasi di Daerah Istimewa Yogyakarta. Penelitian memanfaatkan data penjualan kendaraan bermotor dan data kecelakaan yang diperoleh dari Direktorat Lalu Lintas POLDA DIY. Perumusan masalah pada penelitian ini adalah.

- a. Bagaimana memproses data kecelakaan lalu lintas sehingga dapat mengetahui pola kecelakaan lalulintas berdasarkan merek dan tipe kendaraan.
- b. Bagaimana merumuskan suatu pengetahuan yang berguna dari pola tersebut

Tujuan penelitian ini adalah sebagai berikut :

- a. Melakukan pemrosesan Data Mining Association pada data kecelakaan lalu lintas sehingga kepolisian dapat melihat pola kecelakaan yang ada dan dapat memberikan pembinaan dan penerangan bagi masyarakat maupun dealer kendaraan bermotor.

Metode apriori digunakan untuk memperoleh kaidah asosiasi yang menggambarkan hubungan antar item pada database transaksional. Dari hasil pengujian empiris dapat ditarik kesimpulan bahwa waktu komputasi untuk menghasilkan kaidah asosiasi dipengaruhi oleh jumlah transaksi [5].

Analisis terhadap data kecelakaan lalu lintas. Penelitian ini mengambil lokasi ruas jalan Slamet Riyadi-PB.Soedirman-Sultan Agung-Gajah Mada-Hayam Wuruk di Kabupaten Jember. Dari analisa didapatkan hasil bahwa blackspot dari lokasi studi kecelakaan didapat di jalan Gajah Mada [8].

Data mining merupakan metode untuk menemukan suatu pengetahuan dalam suatu database yang cukup besar. Data mining adalah proses menggali dan menganalisa sejumlah data yang sangat besar untuk memperoleh sesuatu yang benar, baru, sangat bermanfaat dan akhirnya dapat menemukan suatu corak atau pola dalam data tersebut [9].

Data mining adalah bagian integral dari penemuan pengetahuan dalam database yang merupakan proses dengan urutan sebagai berikut.

1. Data cleaning (untuk menghilangkan noise dan inkonsistensi data)
2. Data integration (dimana beberapa data sources akan dikombinasikan)
3. Data selection (dimana hanya data yang dapat dipakai untuk analisis saja yang akan diambil dari database)
4. Data transformation (data akan ditransformasikan ke bentuk yang lebih terstruktur untuk mempermudah proses data mining)
5. Data mining (proses utama data mining dimana teknik data mining diterapkan untuk menghasilkan pola data)
6. Pattern evaluation
7. Knowledge presentation (dimana visualisasi dan representasi hasil diberikan kepada pengguna)

Association rule mining merupakan bagian dari frequent pattern mining. Frequent pattern mining merupakan salah satu [task data mining](#) yang sangat penting. Task ini mencari hubungan/relasi, asosiasi, dan korelasi dalam data. Misalkan terdapat data transaksi seperti pada Tabel 1.

Tabel 1. Contoh data transaksi

<i>TID</i>	<i>Items</i>
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

Contoh pengetahuan yang dapat diperoleh dari data di tabel 3.1. adalah {Beer} → {Diaper}. Artinya orang yang beli beer biasanya beli diaper juga. Algoritma apriori digunakan untuk menghasilkan aturan asosiasi, dengan pola “if-then”. Support (dukungan) adalah probabilitas pelanggan membeli beberapa produk secara bersamaan dari seluruh transaksi. Ukuran support menunjukkan persentase dari keranjang belanja di mana left hand dan right hand keduanya ditemukan secara bersamaan. Support merupakan persentase dukungan pembeli yang membeli barang condition akan membeli juga barang result dari seluruh transaksi yang terjadi :

Contoh:

Transaksi 1	:	A, B, D, E
Transaksi 2	:	B, E, F
Transaksi 3	:	A, D, F
Transaksi 4	:	A, B, C, D, E, F

Tabel 2. Market basket analysis

	A	B	C	D	E	F
A	3	2	1	3	2	2
B	2	3	1	2	3	2
C	1	1	1	1	1	1
D	3	2	1	3	2	2
E	2	3	1	2	3	2
F	2	2	1	2	2	3

Jumlah transaksi dalam Tabel 2. adalah 4, jika seseorang membeli barang D maka akan membeli barang E dengan support $\frac{2}{4} * 100 \% = 50 \%$. support dapat digunakan pada penjualan single item.

Contoh: support(A) = $\frac{3}{4} * 100 \% = 75 \%$.

Baca secara fonetik

2. METODE PENELITIAN

Analisa data mining terhadap data kendaraan bermotor dan data kecelakaan diharapkan dapat memenuhi beberapa kebutuhan yaitu mengemukakan pola kecelakaan kendaraan bermotor berdasarkan pola merek, jenis kendaraan, dan warna.

Analisa pemecahan masalah pola kecelakaan.

1. Data kecelakaan kendaraan bermotor

Data kecelakaan kendaraan bermotor dalam penelitian ini adalah data kecelakaan di tahun 2010. Data kecelakaan terdiri atas dua bentuk yaitu data master kecelakaan mencatat tanggal kejadian, lokasi kejadian kecelakaan, jumlah korban luka-luka dan korban tewas. Bentuk data kecelakaan ada dalam Tabel 3.

Tabel 3. Contoh data master kejadian kecelakaan

ID	TANGGAL	Lokasi	Luka	Tewas
1	1/2/2010	Bantul	1	0
2	5/2/2010	Bantul	1	0
3	8/2/2010	Bantul	2	0
4	17/2/2010	Bantul	1	0

Data detail kecelakaan menjelaskan kendaraan yang terlibat dalam kecelakaan sebagaimana dalam Tabel 4.

Tabel 4. Contoh Data Detil Kejadian Kecelakaan

No	Id	Nopol	Merek
1	1	AB9999CD	YAMAHA
2	1	AB9998CD	HONDA
3	1	AB9997CD	SUZUKI
4	2	AB9996CD	KAWASAKI
5	2	AB9995CD	HONDA
6	2	AB9994CD	TOYOTA
7	3	AB9993CD	YAMAHA
8	3	AB9992CD	KAWASAKI
9	3	AB9991CD	HONDA
10	3	AB9989CD	TOYOTA
11	4	AB9988CD	YAMAHA
12	4	AB9989CD	TOYOTA

Tabel master dan detil diatas berelasi dengan aturan one to many (satu ke banyak).

2. Proses transformasi data

Data kejadian kecelakaan ditransformasikan menjadi bentuk yang mudah diolah oleh proses data mining. Transformasi merek mengikuti kode dalam Tabel 5.

Tabel 5. Kode merek dan merek kendaraan bermotor

KODEMEREK	MEREK
1	YAMAHA
2	HONDA
3	SUZUKI
4	KAWASAKI
5	TOYOTA

Hasil transformasi data kecelakaan ada dalam Tabel 6.

Tabel 6. Hasil transformasi kejadian kecelakaan berdasarkan merek

ID	Merek Terlibat
1	1 3 4
2	2 3 5
3	1 2 3 5
4	2 5

3. Proses data mining apriori

Berikut proses apriori bagi data kecelakaan berdasarkan merek pada Tabel 4. Pada ilustrasi ini ditentukan minimum support adalah 50%. Jumlah kejadian adalah 4 kejadian kecelakaan, maka jika minimum support ditentukan 50% maka minimal jumlah support adalah 2 kejadian. Maka hasil perhitungan apriori untuk itemset-1, 2 dan 3 ada pada Tabel 7.

Tabel 7. Pola Kecelakaan Lalu Lintas

Itemset	Merek	Jumlah Kejadian (support)
{2}	Kawasaki	3
{3}	Honda	3
{5}	Toyota	3
{1}	Yamaha	2
{1 3}	Yamaha, Honda	2
{2 3}	Kawasaki, Honda	2
{2 5}	Kawasaki, Toyota	3
{3 5}	Honda, Toyota	2
{2 3 5}	Kawasaki, Honda, Toyota	2

Pola kecelakaan lalu lintas dapat ditampilkan per wilayah penelitian yaitu untuk Kabupaten Sleman, Bantul, Gunung Kidul, Kulon Progo, atau Kota Yogyakarta. Selain itu juga dapat ditampilkan untuk seluruh wilayah DIY.

3. HASIL DAN PEMBAHASAN

Pola kecelakaan yang dihasilkan adalah jumlah kejadian kecelakaan berdasarkan merek/warna/tipe kendaraan tertentu sehingga dapat diketahui kecenderungan suatu merek/tipe/warna terhadap kejadian kecelakaan.

1. Pola kecelakaan untuk wilayah DIY berdasarkan merek

Pola kecelakaan berdasarkan merek hanya memperhatikan hubungan antara merek dengan kejadian kecelakaan. Tabel 8 memperlihatkan jumlah kejadian kecelakaan bagi merek-merek kendaraan bermotor menggunakan minimum support 10%.

Tabel 8. Pola kecelakaan berdasarkan merek

No.	Merek Kendaraan	Jumlah Kejadian
1	SUZUKI	1695
2	YAMAHA	984
3	HONDA	880

4	KAWASAKI	822
5	HONDA,SUZUKI	326
6	YAMAHA,SUZUKI	322
7	SUZUKI, KAWASAKI	299

2. Pola Kecelakaan untuk wilayah DIY berdasarkan tipe kendaraan

Pola kecelakaan berdasarkan tipe hanya memperhatikan hubungan antara tipe kendaraan dengan kejadian kecelakaan. Tabel 9 memperlihatkan jumlah kejadian kecelakaan berdasarkan tipe kendaraan menggunakan minimum support 5%.

Tabel 9. Pola kecelakaan berdasarkan tipe kendaraan bermotor

No	Tipe Kendaraan	Jumlah Kejadian	Persentase
1.	SEPEDA MOTOR	2692	95.46
2.	SEDAN	560	19.86
3.	MINIBUS	271	9.61
4.	MINIBUS,SEPEDA MOTOR	209	7.41
5.	SEPEDA	183	6.49
6.	SEPEDA MOTOR, SEPEDA	168	5.96
7.	TRONTON	162	5.74
8.	TIDAK DIKETAHUI	160	5.67

4. KESIMPULAN

Setelah proses penelitian dilakukan diperoleh beberapa kesimpulan tentang hasil data mining terhadap data penjualan kendaraan bermotor dan data kecelakaan sebagai berikut :

1. Pola kecelakaan telah diketahui. Kecelakaan sering terjadi untuk kendaraan dengan merek Suzuki, disusul Yamaha dan Kawasaki.
2. Jika digunakan kriteria jenis kendaraan, kecelakaan sering terjadi pada sepeda motor, disusul dengan sedan dan kemudian minibus.
3. Pola kecelakaan berdasarkan warna kendaraan tidak dapat dilakukan karena data warna kendaraan yang terlibat kecelakaan tidak dicatat dalam laporan polisi.

5. SARAN

Penelitian ini memiliki beberapa kelemahan sehingga membutuhkan saran-saran untuk pengembangan penelitian ini di masa depan.

1. Sebagai saran untuk DITLANTAS POLDA DIY adalah hendaknya proses pemasukan data menggunakan form standar dimana pengisian nama dirinci menjadi nama depan, tengah, dan belakang, kemudian pengisian alamat juga dikodekan berdasarkan nomor RT, RW, Kelurahan, Kecamatan, dan Kabupaten, atau dengan kode nama jalan yang seragam. Input data secara benar akan mempermudah proses data cleaning.
2. Pola kecelakaan hanya didasarkan pada data yang ada di Kepolisian. Kejadian kecelakaan yang tidak dilaporkan tidak dapat diproses. Hendaknya pihak Kepolisian dapat merekam kejadian kecelakaan semaksimal mungkin dengan melibatkan masyarakat.
3. Penelitian perlu dikembangkan untuk metode yang lain, misalnya dengan klasifikasi jenis lain atau menggunakan metode yang lebih baik.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ditlantas POLDA DIY yang telah mengijinkan peneliti memanfaatkan data kendaraan bermotor dari bagian BPKB dan data kecelakaan. Semoga penelitian ini bermanfaat bagi semua pihak.

DAFTAR PUSTAKA

- Dezhi, X., dan Ganegoda, G.U, 2011, Rule Based Classification to Detect Malnutrition in Children, International Journal on Computer Science and Engineering (IJCSE) Vol. 3 No. 1 Jan 2011 ISSN : 0975-3397
- Firmansyah, A., 2009, Analisa Kecelakaan Lalu Lintas Kabupaten Jember (Ruas Jalan Slamet Riyadi - Pb.Sudirman - Sultan Agung – Gajah Mada - Hayam Wuruk), Tugas Akhir Jurusan Teknik Sipil Fakultas Teknik Sipil Dan Perencanaan Institut Teknologi Sepuluh Nopember Surabaya
- Han, J., and Kamber, M., 2006, Data Mining Concept and Technique, Morgan Kaufman Publisher, San Francisco
- Marthin, J. dan Samuel, H., 2007, Analisis Tingkat Brand Loyalty Pada Produk Shampoo merek "Head & Shoulders", Jurnal Manajemen Pemasaran, Vol. 2, No. 2, Oktober 2007: 90-102, Universitas Kristen Petra, Surabaya
- Muchsin, N., Kiptiyah, S.M., Martawidjaja, S., 2003, Analisis Sikap Konsumen Dalam Keputusan Pembelian Sepeda Motor Merek Sanex Dan Kanzen Di Kota Malang, Jurnal Fakultas Ekonomi Universitas Brawijaya Malang
- Murugan, S., 2010, Rule Based Classification Of Ischemic Ecg Beats Using Antminer, International Journal of Engineering Science and Technology Vol. 2(8), 2010, 3929-3935
- Peter, J. P and Olson J. C. 2001. Consumer Behavior and Market Strategy, McGraw-Hill College, New York
- Santoso, L.W., 2009, Pembuatan Perangkat Lunak Data Mining Untuk Penggalan Kaidah Asosiasi Menggunakan Metode Apriori, Jurnal Informatika Fakultas Teknologi Industri, Jurusan Teknik Informatika, Universitas Kristen Petra
- Tada, T., Nagashima, T., Okada, Y., 2010, Rule-based Classification for Audio Data Based on Closed Itemset Mining, Proceedings of the International MultiConference of Engineers and Computer Scientist 2010 Vol I, Hongkong
- Qin, B., dan Xia, Y., 2009, A Rule-Based Classification Algorithm for Uncertain Data, IEEE International Conference on Data Engineering
-