

Data Mining Untuk Mengetahui Tingkat Loyalitas Konsumen Terhadap Merek Kendaraan Bermotor dan Pola Kecelakaan Lalulintas di DIY

A. Sasmito Aribowo dan E. Winarko

Abstract— *The data of vehicle sales and traffic accident can be processed into information that is important for vehicle dealers and the Police Department. Those important information researched are the level of consumer loyalty to the vehicle brands and to predict the vehicle's brands that will be purchased by a consumer. The study also tries 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 called 'rule based classification' to establish the sales of vehicles rules by which can be used to classify consumer into group level of brand loyalty and also estimate the brand of the next vehicle's brand that will be purchased by the consumer. This research will 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 companies or vehicle dealers to obtain information about the level of the consumer's brand loyalty to the dealer's brand and to predict the brand that the consumer would be buy for the next vehicle. The result can also 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.*

1. PENDAHULUAN

Salah satu bentuk interaksi yang bersifat kausal dan saling membutuhkan adalah interaksi antara dealer penjualan kendaraan bermotor dengan Kepolisian Republik Indonesia (POLRI).

Agus Sasmito Ariwibowo Mahasiswa Program Pascasarjana Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada, Yogyakarta.

Edi Winarko, Staff Pengajar Program Pascasarjana Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada, Yogyakarta.

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 dealer kendaraan bermotor adalah informasi tentang tingkat kesetiaan konsumen terhadap merek kendaraan yang dijualnya, serta bagaimana memprediksi merek kendaraan yang akan dibeli oleh seorang konsumen.

Di sisi lain kepolisian memiliki data 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.

- Bagaimana menghitung tingkat *brand loyalty* konsumen untuk setiap merek kendaran bermotor.
- Bagaimana membantu dealer penjual kendaraan bermotor memprediksi merek kendaraan bermotor apakah yang akan dibeli berikutnya oleh seorang konsumen berdasarkan tingkat *brand loyalty* seorang.
- Bagaimana memproses data kecelakaan lalu lintas sehingga dapat mengetahui pola kecelakaan lalulintas berdasarkan merek, tipe dan warna kendaraan.

Tujuan penelitian ini adalah sebagai berikut :

- Analisis terhadap data kendaraan bermotor dan mengubahnya menjadi data pola merek kendaraan yang dibeli oleh setiap konsumen.

- b. Menerapkan teknik data mining *rule based classification* untuk mengklasifikasikan konsumen dalam kelompok-kelompok konsumen dengan tingkatan loyalitas merek tertentu sehingga dapat dipakai juga untuk memperkirakan merek kendaraan berikutnya yang akan dibeli oleh seorang konsumen.
- c. 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.

Penelitian tentang metode dan domain yang sama pernah dilakukan sebelumnya. *Rule based classification* dapat dipakai untuk memproses ketidakpastian antara data numerik dan kategoris. Hasil penelitian menunjukkan bahwa *rule based classification* yang memiliki performa yang sangat baik bahkan ketika data sangat tidak pasti [1].

Rule based classification dapat dipakai untuk melakukan klasifikasi data audio multi-kelas [2]. *Rule based classification* pernah dipakai untuk membandingkan secara akurat denyut jantung orang sakit dengan denyut iskemik yang sehat [3]. *Ruled based classification* juga dapat digunakan untuk mendeteksi kekurangan gizi pada anak-anak [4].

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].

Penelitian tentang loyalitas merek dan bertujuan untuk mengetahui tingkat *brand loyalty* konsumen atas *shampoo* merek *Head & Shoulders* pernah dilakukan. Penelitian ini juga hendak mengetahui susunan piramida loyalitas, yang meliputi *switcher*, *habitual buyer*, *satisfied buyer*, *liking of the brand*, dan *committed buyer* atas *shampoo* merek *Head & Shoulders*. Hasil penelitian atas 1200 orang pelanggan menunjukkan bahwa *Shampoo* merek *Head & Shoulders* mempunyai konsumen yang paling banyak pada tingkat *committed buyer*, yaitu 91,25%, sehingga disimpulkan bagus [6].

Penelitian untuk mengetahui dan menganalisis sikap konsumen dalam pembelian sepeda motor merek Sanex dan Kanzen pernah dilakukan. Penelitian dilakukan di Kota Malang. Hasil pengujian yang diperoleh menunjukkan bahwa

rata-rata kedua kelompok konsumen ini memiliki kesamaan sikap dalam pembelian sepeda motor merek Sanex atau Kanzen [7].

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.

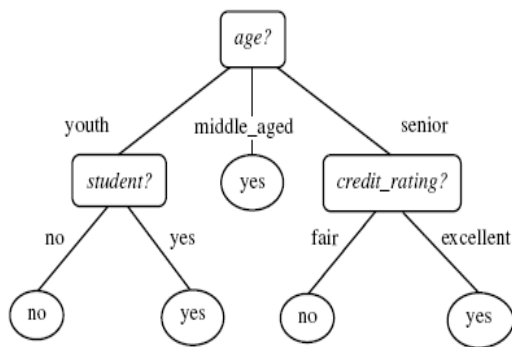
1. *Data cleaning* (untuk menghilangkan noise dan inkonsistensi data)
2. *Data integration* (dimana beberapa data sources akan dikombinasikan)
3. *Data selection* (dimana hanya data yang dapat dipakai untuk analisis saja yang akan diambil dari database)
4. *Data transformation* (data akan ditransformasikan 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)

Tujuan utama penerapan *data mining* adalah untuk prediksi (*prediction*) dan uraian (*description*).

Klasifikasi adalah proses menemukan model (atau fungsi) yang akan mengelompokkan kelas data sehingga dapat memprediksi kelas obyek yang tidak diketahui. Model yang diperoleh bisa diwakili di berbagai bentuk, seperti klasifikasi berbentuk aturan IF-THEN, pohon keputusan, rumus matematika, atau jaringan saraf tiruan [9].

Rule based classification adalah model berbentuk seperangkat aturan berbentuk IF-THEN. Aturan berbentuk IF-THEN adalah cara yang baik untuk merepresentasikan informasi atau pengetahuan. Aturan dapat diekstraksi dari pohon keputusan. Untuk mengekstrak aturan dari

sebuah pohon keputusan maka caranya adalah dengan membuat satu aturan untuk setiap garis yang menghubungkan akar ke simpul daun. Setiap menjumpai sebuah kriteria dalam menelusuri garis tersebut maka aturan perlu diberi operator logika AND untuk membentuk aturan bagian "JIKA" (anteseden). Simpul daun yang menunjukkan sebuah kelas, membentuk konsekuen atau bagian "THEN".



Gambar 1. Contoh Pohon Keputusan [9]

Pohon Keputusan pada gambar 1 dapat dikonversi untuk klasifikasi aturan IF-THEN dengan menelusuri jalur dari *node root* setiap *node* daun di pohon. Aturan diekstrak dari gambar 1 adalah :

- R1: JIKA age = youth DAN student = no MAKA membeli_komputer = no
- R2: JIKA age = youth DAN student = yes MAKA membeli_komputer = yes
- R3: JIKA age = middle_aged MAKA membeli_komputer = yes
- R4: JIKA age = senior DAN credit_rating = excellent MAKA membeli_komputer = yes
- R5: JIKA age = senior DAN credit_rating = fair THEN membeli_komputer = no

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, assosiasi, dan korelasi dalam data. Misalkan terdapat data transaksi seperti pada Tabel 1.

Tabel 1. Contoh data transaksi

TID	Items
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 $2/4 * 100 \% = 50 \%$. support dapat digunakan pada penjualan single item.

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

Loyalitas merek sudah lama menjadi gagasan sentral dalam pemasaran, merupakan satu ukuran keterkaitan seorang pelanggan pada merek. Pola pembelian dan Brand Purchase ada dalam beberapa tingkatan sebagaimana dalam Tabel 3 [10].

Tabel 3. Purchase pattern categories and brand purchase sequence [10]

Purchase Pattern Category	Brand Purchase Sequence
Undevided Brand Loyalty (UBL)	AAAAAAAAAA
Occasional Swith Brand Loyalty (OS)	AAABCAAD
Swith Brand Loyalty (SBL)	AAAAABBBBB
Devided Brand Loyalty (DBL)	AABABBAABB
Brand Indifference(BI)	ABCDEFGHIJ

- Undevided brand loyalty :

Merupakan kondisi ideal, konsumen hanya mengkonsumsi satu merek dan terus melakukan pembelian ulang.

- Brand loyalty with occasional switch

Pada saat tertentu kemungkinan konsumen berpindah ke merek lain dengan berbagai alasan, misalnya : tidak ada stock persediaan, competitor menawarkan harga special, mencoba produk baru yang ada di pasaran.

- *Brand loyalty swiches* :
Merupakan konsumen yang berpindah dari sebuah merek ke merek lain dengan berbagai alasan.
- *Devided brand loyalty* :
Pembelian dua merek atau lebih secara konsisten. Misalnya dalam sebuah keluarga menggunakan tiga jenis shampo yang berbeda. Satu untuk orang tua, satu untuk anak remaja, satu untuk adik bayi.
- *Brand indifference* :
Pembelian merek tanpa pola pembelian ulang. Kebalikan dari *undevided brand loyalty*.

2. METODE PENELITIAN

Analisa data mining terhadap data kendaraan bermotor dan data kecelakaan diharapkan dapat memenuhi beberapa kebutuhan berikut.

1. Mampu menggali informasi tingkat loyalitas merek konsumen untuk setiap merek kendaraan bermotor. Loyalitas dibagi menjadi dua yaitu loyalitas pembelian sepeda motor dan kendaraan non sepeda motor)
2. Mampu memprediksi merek kendaraan yang hendak dibeli berikutnya berdasarkan loyalitas merek seorang konsumen.
3. Mampu mengemukakan pola kecelakaan kendaraan bermotor berdasarkan pola merek, jenis kendaraan, dan warna.

A. Analisa pemecahan masalah loyalitas merek.

Proses pemenuhan kebutuhan untuk permasalahan pada nomor 1 dan 2 dilakukan dengan langkah-langkah sebagai berikut :

1. Mendapatkan data penjualan kendaraan bermotor.
2. Melakukan *Data Cleaning*
3. Melakukan Transformasi Data.
4. Membangun *Rule Based Classifier*
5. Melakukan Ekstraksi Ciri dan Klasifikasi Data Penjualan Kendaraan Bermotor.
6. Menampilkan Hasil Klasifikasi Per Merek Kendaraan
7. Melakukan Prediksi Pembelian Berikutnya Untuk Setiap Konsumen

Secara mendetil penjelasan proses setiap langkah diatas adalah sebagai berikut.

1. Proses transformasi data.

Data penjualan kendaraan ditransformasikan menjadi data *sequence* (urutan) merek kendaraan bermotor yang dibeli oleh seorang konsumen sebagaimana pada Tabel 4.

Tabel 4. Hasil transformasi *sequence* data pembelian per konsumen

IDPEMBELI	SEQUENCE MEREK
1	AAA
2	AAB
3	ABAB
4	ABCD

2. Proses pembuatan *rule based classifier*

Klasifikasi tingkat loyalitas konsumen dibagi menjadi lima tingkatan yaitu Undevided Brand Loyalty, Brand Loyalty with Occasional Switch, Brand Loyalty Switches, Devided Brand Loyalty, dan Brand Indifference (Peter & Olson, 2001). Bentuk *sequence* merek dalam pembelian untuk setiap jenis tingkatan loyalitas konsumen ada dalam Tabel 5.

Sebelum proses klasifikasi, untuk mempermudah pengelompokan konsumen pada loyalitas merek perlu dilakukan pembentukan aturan klasifikasi (*rule base classifier*). Aturan tersebut digunakan untuk ekstraksi ciri *sequence* pembelian merek setiap konsumen. Aturan (*rule*) untuk klasifikasi dibentuk dengan menganalisa pola *sequence* pembelian setiap jenis loyalitas merek dalam Tabel 5.

Tabel 5. Klasifikasi tingkat loyalitas konsumen

No.	Jenis	Bentuk Sequence Merk
1.	<i>Undevided Brand Loyalty</i>	AAAAAAA
2.	<i>Occasional Switch</i>	AABAACAD
3.	<i>Switch</i>	AAAABBB
4.	<i>Devided Brand Loyalty</i>	AABBAABA
5.	<i>Brand Indifference</i>	ABCDEFGF

Beberapa hal yang diamati pada pola merek yang dibeli oleh konsumen pada Tabel 5 adalah :

- a. Jumlah jenis merek dalam *sequence* pembelian seorang konsumen.
- b. Jumlah *switch* (perpindahan) dari satu merek ke merek lain dalam *sequence* pembelian seorang konsumen.
- c. Apakah ada *switch* kembali ke merek awal yang dipakai sebagai acuan loyalitas.

Hasil ekstraksi ciri akan dijadikan aturan yang akan dipakai sebagai pedoman klasifikasi. Hasil ekstraksi ciri tersebut ada di dalam Tabel 6.

Tabel 6. Ekstraksi ciri pada tingkat loyalitas konsumen

Jenis Brand Loyalty	Bentuk Sequence Pembelian	EKSTRAKSI CIRI		
		Cacah Jenis Merek	Cacah Switch Merek	Ada Beli Kembali Ke Merk Awal
Undevided	AAAAA	1	-	Tidak
Occasional Switch	ABACAD	> 2	> 1	Ya
Switch	AAABB	2	1	Tidak
Devided	AABABA	2	> 1	Ya
Brand Indifference	ABCDEF	>2	=Cacah Jenis-1	Tidak

Keterangan :

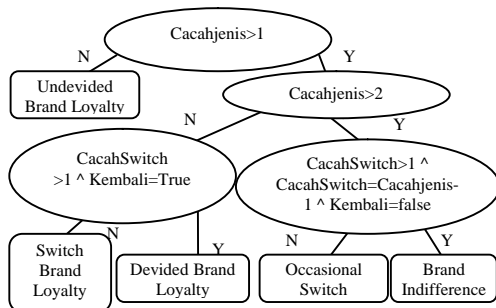
Cacahjenis : cacah jenis merek yang dibeli konsumen

CacahSwitch : cacah perpindahan merek dibeli konsumen

Kembali : ada pembelian kembali merek awal (true/false)

3. Pembentukan aturan klasifikasi (*Rule Based Classifier*)

Tabel 6 pada kolom ekstraksi ciri diubah menjadi aturan klasifikasi (*rule base classifier*). *Rule base* disusun menggunakan aturan pembentukan *rule base* yang dikenal dengan pendekatan *direct method* atau pendekatan langsung. Pembentukan *rule base* diawali dengan membuat pohon keputusan sebagaimana pada Gambar 2. yang akan mempermudah dalam menurunkan *rule* dari Tabel 6.



Gambar 2. Pohon Keputusan Untuk Proses Klasifikasi

4. Rule set

Pohon keputusan tersebut dapat dibentuk *rule set* yang berguna untuk proses klasifikasi. Sebuah *rule* dalam *ruleset* memiliki *anteseden* dan *konsequen*. Sebagai *anteseden* adalah ciri dari setiap *brand loyalty*. Kemudian sebagai *konsequen* adalah nama *brand loyalty*. *Rule set* yang dapat dihasilkan dari pohon keputusan pada Gambar 2 adalah sebagai berikut :

R1 : IF Cacahjenis>1=NO THEN Undevided Brand Loyalty

R2 :IF Cacahjenis>1=YES AND Cacahjenis>2=NO AND

Cacahswitch>1=NO AND Kembali=False THEN Switch Brand Loyalty

R3 : IF Cacahjenis>1=YES AND Cacahjenis>2=NO AND Cacahswitch>1=YES AND Kembali=True THEN Devided Brand Loyalty

R4 : IF Cacahjenis>1=YES AND Cacahjenis>2=YES AND ((CacahSwitch>1) ^ (CacahSwitch=Cacahjenis-1))=NO AND Kembali=True THEN Occasional Switch

R5 : IF Cacahjenis>1=YES AND Cacahjenis>2=YES AND ((CacahSwitch>1) ^ (CacahSwitch=Cacahjenis-1))=NO AND Kembali=False THEN Brand Indifference

5. Proses ekstraksi ciri dan klasifikasi data pembelian konsumen

Setelah *rule* klasifikasi terstruktur dengan baik maka *rule* tersebut telah siap digunakan untuk mengklasifikasikan loyalitas merek semua konsumen kendaraan bermotor. Metode ini disebut sebagai *rule based classification data mining* yang diterapkan pada *sequential data*. Setiap konsumen akan diteliti satu per satu tentang tiga hal berikut ini :

- Jumlah jenis merek dalam *sequence* pembelian seorang konsumen
- Jumlah *switch* (perpindahan) merek dalam *sequence* pembelian konsumen.
- Apakah ada *switch* kembali ke merek awal sebagai merek acuan loyalitas.

Tabel 7 pada kolom kelompok dapat diisi nama *brand loyalty* sesuai ciri yang diperoleh (untuk setiap konsumen).

Tabel 7. Hasil Proses Ekstraksi Ciri dan Pencocokan Pola Penjualan pada Sequence Pembelian Setiap Konsumen

Id Pembeli	Urutan Merek	Cacah Switch	Cacah Merek	Kembali (1=Ya, 0=Tidak)	Loyalitas Merek
1	AAA	0	1	1	UBL
2	AAB	1	2	0	SBL
3	ABAB	3	2	1	DBL
4	ABCD	3	4	0	BI

6. Hasil klasifikasi per merek kendaraan

Hasil klasifikasi per merek kendaraan merupakan hasil pencacahan jumlah setiap jenis loyalitas merek kendaraan bermotor untuk merek tertentu sebagaimana dalam Tabel 8.

Tabel 8. Hasil klasifikasi loyalitas merek untuk kendaraan bermerek X

Jenis	Jumlah	Persentase
UBL	33315	61.13
OS	273	0.5
SBL	16402	30.09
DBL	2792	5.12
BI	1719	3.15

7. Proses prediksi pembelian berikutnya untuk setiap konsumen

Setelah jenis loyalitas merek setiap konsumen diketahui, maka proses prediksi pembelian berikutnya dapat dilakukan. Prediksi dilakukan dengan melakukan analisa prediksi setiap pola pembelian dalam Tabel 9.

Berdasarkan aturan analisa prediksi pembelian berikutnya pada Tabel 9, maka setiap konsumen dalam Tabel 7 dapat dihitung dan diketahui prediksi pembelian berikutnya. Hasil prediksi pembelian berikutnya ada dalam Tabel 10.

Tabel 9. Analisa Prediksi Pembelian Berikutnya untuk Setiap Pola Klasifikasi

Jenis	Pola Pembelian (Merek)	Prediksi Pembelian Berikutnya
UBL	AAAAA	A=100%, Selain A = 0 %
OS	ABACAD	A=50%, Selain A = (1/Jumlah Merk Selain A)*50%
SBL	AABBB	A=0%, B= 100%
DBL	ABABA	A=50%, B=50%
BI	ABCDE	Merek selain ABCDE

Tabel 10. Analisa prediksi pembelian berikutnya untuk setiap konsumen

Id Pembeli	Sequence Pattern (Merek)	Prediksi Pembelian Berikutnya
1	AAA	A
2	AAB	B
3	ABAB	A atau B
4	ABCD	SELAIN ABCD

B. 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 11.

Tabel 11. 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 detil kecelakaan menjelaskan kendaraan yang terlibat dalam kecelakaan sebagaimana dalam Tabel 12.

Tabel 12. 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 13.

Tabel 13. 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 14.

Tabel 14. 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 14. 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 15.

Tabel 15. 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

A. Loyalitas Merek

Setelah proses data mining diperoleh hasil tentang loyalitas umum semua pengguna (konsumen) kendaraan bermotor di Daerah Istimewa Yogyakarta. Hasil loyalitas umum ada di Tabel 16.

Tabel 16. Loyalitas konsumen terhadap merek kendaraan bermotor di DIY

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	12.734	7,92
2	DBL	9.840	6,12
3	OS	1.205	0,75
4	SBL	74.512	46,34
5	UBL	62.514	38,88
	Total	160.805	100

Tabel 16 menunjukkan bahwa konsumen di DIY memiliki sifat dominan untuk berpindah merek dari satu merek ke merek yang lain dengan jumlah konsumen yang berpindah adalah 46,34%. Tabel tersebut juga menunjukkan bahwa jumlah konsumen yang setia terhadap suatu merek cukup besar yaitu sebesar 38,88%.

1. Loyalitas kendaraan sepeda motor

Sepeda motor Honda menempati urutan pertama yang memiliki jumlah konsumen loyal terbanyak. Hasil perhitungan ada di Tabel 17. Sebagai pembanding adalah merek competitor terkuat Honda yaitu Yamaha, Suzuki dan Kawasaki. Masing-masing ada di Tabel 18, Tabel 19, dan Tabel 20.

Tabel 17. Loyalitas konsumen sepeda motor Honda

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	3.118	4.0447
2	DBL	5.435	7.0503
3	OS	571	0.7407
4	SBL	24.751	32.1070
5	UBL	42.994	55.7719
	Jumlah	76.869	100

Tabel 18. Loyalitas konsumen terhadap sepeda motor merek Yamaha

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	3.114	8.4185
2	DBL	2.253	6.0908
3	OS	274	0.7407
4	SBL	20.818	56.2801
5	UBL	10.433	28.2049
	Jumlah	36.892	100

Tabel 19. Loyalitas konsumen terhadap sepeda motor merek Suzuki

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	3.320	10.0387
2	DBL	1.877	5.6755
3	OS	271	0.8194
4	SBL	19.391	58.6327
5	UBL	8.089	24.4588
	Jumlah	32.948	100

Tabel 20. Loyalitas konsumen terhadap sepeda motor merek Kawasaki di DIY

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	624	17.8184
2	DBL	81	2.3130
3	OS	29	0.8281
4	SBL	2.405	68.6750
5	UBL	348	9.9372
	Jumlah	3.487	100

Dari Tabel 17 hingga Tabel 20. dapat disimpulkan bahwa kendaraan sepeda motor yang memiliki jumlah konsumen dengan loyalitas setia terbesar secara berurutan adalah Honda (55,68%), Yamaha (28,22%), Suzuki (24,5%), dan Kawasaki (9,96%).

2. Loyalitas kendaraan non sepeda motor (mobil)

Mobil Toyota menempati urutan pertama yang memiliki jumlah konsumen loyal terbanyak. Hasil perhitungan ada di Tabel 22. Sebagai pembanding adalah merek Honda, Suzuki dan Daihatsu. Masing-masing ada di Tabel 21, Tabel 23, dan Tabel 24.

Tabel 21. Loyalitas konsumen terhadap mobil merek Honda

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	272	15.32
2	DBL	83	4.67
3	OS	21	1.18
4	SBL	1.009	56.85
5	UBL	383	21.58
	Jumlah	1.768	100

Tabel 22. Loyalitas konsumen terhadap mobil Toyota

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	469	10.2986
2	DBL	248	5.4458
3	OS	36	0.7905
4	SBL	2.171	47.6724
5	UBL	1.575	34.5850
	Jumlah	4.499	100

Tabel 23. Loyalitas konsumen terhadap mobil Suzuki

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	375	11.3533
2	DBL	134	4.0569
3	OS	29	0.8780
4	SBL	1.759	53.2546
5	UBL	991	30.0030
	Jumlah	3.288	100

Tabel 24. Loyalitas konsumen terhadap mobil Daihatsu

No	Brand Loyalty	Jumlah Konsumen	(%)
1	BI	262	11.4812
2	DBL	116	5.0833
3	OS	25	1.0955
4	SBL	1.300	56.9676
5	UBL	566	24.8028
	Jumlah	2.269	100

Dari Tabel 21 hingga Tabel 24 dapat disimpulkan bahwa kendaraan mobil yang memiliki jumlah konsumen dengan loyalitas setia terbesar secara berurutan adalah Toyota (34,585%), Suzuki (30.00%), Daihatsu (24,8%), dan Honda (21,58%).

B. Pola kecelakaan

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 25 memperlihatkan jumlah kejadian kecelakaan bagi merek-merek kendaraan bermotor menggunakan minimum support 10%.

Tabel 25. 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 26 memperlihatkan jumlah kejadian kecelakaan berdasarkan tipe kendaraan menggunakan minimum support 5%.

Tabel 26. Pola kecelakaan berdasarkan tipe kendaraan bermotor

No	Tipe Kendaraan	Jumlah Kejadian
1.	SEPEDA MOTOR	2692
2.	SEDAN	560
3.	MINIBUS	271
4.	MINIBUS,SEPEDA MOTOR	209
5.	SEPEDA	183
6.	SEPEDA MOTOR, SEPEDA	168
7.	TRONTON	162
8.	TIDAK DIKETAHUI	160

C. Pengujian Hasil Data Mining

1. Hasil klasifikasi loyalitas merek

Pengujian dilakukan dengan mengambil semua hasil proses klasifikasi dengan ketentuan jumlah pembelian minimal 3 dan dilakukan pada konsumen *Undivided Brand Loyalty* (setia) untuk pembelian hingga ke n-1 (n adalah jumlah cacah pembelian). Jumlah data uji adalah sejumlah 13819 konsumen. Kemudian dilakukan pengamatan apakah pembelian ke n (pembelian terakhir) adalah merek yang sama dengan pembelian sebelumnya atau tidak. Jika sama maka konsumen tersebut adalah jenis konsumen yang tetap setia, dan jika tidak maka konsumen tersebut adalah jenis konsumen yang melakukan *Switch* (beralih merek). Berdasarkan hasil

pengujian diperoleh informasi bahwa konsumen setia yang tetap membeli kendaraan yang sama dengan merek pada pembelian sebelumnya adalah sejumlah 9120 orang konsumen dari total 13819 orang konsumen setia. Jumlah ini adalah sekitar 66%. Sedangkan konsumen setia yang beralih ke merek lain adalah sejumlah 4699 konsumen atau sekitar 34%.

2. Pengujian prediksi pembelian berikutnya

Pengujian prediksi pembelian berikutnya dilakukan pada konsumen dengan jenis loyalitas *Undevided Brand Loyalty*, yaitu konsumen yang setia.

a. Pengujian pada konsumen sepeda motor

Hasil perhitungan terhadap total 757 pembelian kembali di tahun 2011 mulai Januari 2011 hingga Maret 2011 bagi konsumen yang pernah membeli di tahun sebelumnya khusus sepeda motor diperoleh hasil dalam Tabel 27.

Tabel 27. Perbandingan jumlah konsumen yang loyal dan konsumen yang beralih ke merek lain di tahun 2011 (khusus sepeda motor)

No	Perilaku Konsumen Loyal	Jumlah	Persentase
1	Beralih ke merek lain	263	34,75 %
2	Tidak beralih merek	494	65,25 %
	Jumlah	757	100%

Berdasarkan hasil proses data mining merek Honda merupakan merek yang paling tinggi tingkat kesetiaan konsumennya. Maka Tabel 28 merupakan tabel khusus untuk membandingkan jumlah konsumen yang beralih merek dari sepeda motor Honda ke merek lain dan konsumen yang tetap setia.

Tabel 28. Perbandingan jumlah konsumen yang loyal dan konsumen yang beralih ke merek lain di tahun 2011 (khusus sepeda motor HONDA)

No	Konsumen Loyal	Jumlah	Persentase
1	Beralih merek	68	16,4 %
2	Tidak beralih merek	349	83,6 %
	Jumlah	417	100%

Berdasarkan hasil pada Tabel 28 dapat diketahui bahwa khusus untuk merek Honda, konsumen loyal yang tetap pada merek Honda adalah sejumlah 83,6%. Merek Yamaha adalah 60,57 %, dan Suzuki 45,6% (tabel tidak disertakan)

b. Pengujian pada konsumen mobil

Hasil perhitungan terhadap total 89 pembelian kembali di awal tahun 2011 mulai Januari 2011

hingga Maret 2011 bagi konsumen yang pernah membeli di tahun sebelumnya khusus mobil diperoleh hasil dalam Tabel 29.

Tabel 29. Perbandingan jumlah konsumen yang loyal dan konsumen yang beralih ke merek lain di Tahun 2011 (khusus mobil)

No	Perilaku Konsumen Loyal	Jumlah	Persentase
1	Beralih ke merek lain	40	44,94 %
2	Tidak beralih merek	49	55,06 %
	Jumlah	89	100%

Mobil merek Toyota merupakan merek yang paling tinggi tingkat kesetiaan konsumennya. Maka Tabel 30 merupakan tabel khusus untuk membandingkan jumlah konsumen yang beralih merek dari mobil Toyota ke merek lain dan konsumen yang tetap setia.

Tabel 30. Perbandingan jumlah konsumen yang loyal dan konsumen yang beralih ke merek lain di tahun 2011 (khusus mobil Toyota)

No	Konsumen Loyal	Jumlah	Persentase
1	Beralih merek	15	34,88%
2	Tidak beralih merek	28	65,12%
	Jumlah	43	100%

Berdasarkan hasil pada Tabel 30 dapat diketahui bahwa khusus untuk merek mobil Toyota, konsumen loyal yang tetap pada merek Toyota adalah sejumlah 65,12%. Merek Suzuki 50%, dan Daihatsu 60% (tabel tidak ditampilkan).

4. KESIMPULAN

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

1. Hasil proses data mining menunjukkan bahwa secara umum tanpa memandang merek kendaraan bermotor, konsumen yang bersifat setia dengan jenis *Undevided Brand Loyalty* adalah sejumlah 38,88%. Konsumen yang melakukan perpindahan merek ke merek yang lain (*Devided Brand Loyalty/Switch*) sejumlah 46,36%. Jenis loyalitas yang lain rata-rata dibawah 10%.
2. Merek sepeda motor dengan tingkat kesetiaan tertinggi (*Undevided Brand Loyalty*) adalah merek Honda (55,68%), diikuti merek Yamaha (28,22%), Suzuki (24,5%), dan Kawasaki (9,96%). Sedangkan mobil adalah Toyota (34,59%), Suzuki (30%), Daihatsu (24,8%) dan Honda (21,58%).

3. Berdasarkan pengamatan pada jumlah persentase konsumen yang tetap loyal pada tahun 2011, penelitian ini membuktikan bahwa sepeda motor Honda dan mobil Toyota merupakan merek yang memberikan kekuatan pemasaran yang lebih baik daripada merek lain.
4. Proses peramalan pembelian berikutnya oleh konsumen sudah dapat dilakukan dengan mengetahui pola loyalitas konsumen yang hendak diramal tersebut.
5. Pola kecelakaan telah diketahui. Kecelakaan sering terjadi untuk kendaraan dengan merek Suzuki, disusul Yamaha dan Kawasaki. Jika digunakan kriteria jenis kendaraan, kecelakaan sering terjadi pada sepeda motor, disusul dengan sedan dan kemudian minibus. 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 mengizinkan peneliti memanfaatkan data kendaraan bermotor dari bagian BPKB dan data kecelakaan. Semoga penelitian ini bermanfaat bagi semua pihak.

DAFTAR PUSTAKA

- [1] Qin, B., dan Xia, Y., 2009, *A Rule-Based Classification Algorithm for Uncertain Data*, IEEE International Conference on Data Engineering
- [2] 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
- [3] 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
- [4] Dezhi, X., dan Ganegoda, G.U., 2011, *Rule Based Classification to Detect Malnutrition in Children*, International Journal on Computer Science and Engineering (IJCSSE) Vol. 3 No. 1 Jan 2011 ISSN : 0975-3397
- [5] 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
- [6] 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
- [7] 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
- [8] 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
- [9] Han, J., and Kamber, M., 2006, *Data Mining Concept and Technique*, Morgan Kaufman Publisher, San Francisco
- [10] Peter, J. P and Olson J. C. 2001. *Consumer Behavior and Market Strategy*, McGraw-Hill College, New York