

Implementasi Data Mining Menggunakan Algoritma Apriori Pada Penjualan Sepeda Motor Jenis Honda (Studi Kasus : Showroom Honda Arista Pematangsiantar)

1) **Lestari Sinaga**

STIKOM Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia
E-Mail: lestarisinaga445@gmail.com

2) **Abdullah Ahmad**

AMIK Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia
E-Mail: abdul@amiktunasbangsa.ac.id

3) **M.Safii**

AMIK Tunas Bangsa Pematangsiantar, Sumatera Utara, Indonesia
E-Mail: m.safii@amiktunasbangsa.ac.id

ABSTRACT

Honda Arista showroom is one of the showrooms or dealers that provide sale and purchase of Honda motorcycle vehicles with various types and prices. Data mining is a series of processes to explore the added value of a data collection in the form of knowledge that has not been known manually. One method that data mining used in this study is the Apriori Method. The Apriori method is one of the data mining techniques that functions to find associative rules between an item combination. The items analyzed are goods ordered together with other items or in other words ordering more than one item but still involving the overall ordering data. The items included in the sales data are Revo, Supra, Beat, Vario, Sonic, PCX, CB, Scoopy, Megapro, Verza, CBR. The results of the association in the form of information about what brands of Honda are purchased simultaneously by consumers, can be used as material for consideration to establish a Marketing Strategy and on transaction data.

Keyword: Honda Showroom, Data Mining, Apriori

PENDAHULUAN

Indonesia terkenal sebagai masyarakat yang memiliki kebiasaan pola hidup konsumtif. Kebiasaan hidup konsumtif tersebut tentunya dengan meningkatnya budaya membeli barang, baik bahan pangan, maupun bahan pakaian, termasuk kendaraan bermotor. Untuk saat ini, sarana transportasi yang sangat bisa diandalkan untuk kalangan kelas menengah yaitu Kenderaan bermotor. Dan kenderaan bermotor yang menjadi andalan untuk saat ini adalah sepedamotor Honda. Keunggulan sepedamotor Honda sudah menjadi satu image tersendiri di kalangan masyarakat karena performa mesin yang canggih dan awet. Data kendaraan bermotor dan pemanfaatan dapat diolah menjadi informasi bagi Dealer penjualan kendaraan bermotor. Showroom Honda Arista merupakan salah satu *showroom* atau *Dealer* yang menyediakan jual beli kendaraan sepeda motor dikota Pematangsiantar.

Honda Arista sendiri menyediakan berbagai jenis sepeda motor Honda dengan berbagai type dan harga. Diantaranya sepeda motor yang cukup banyak diminati dan dibeli pelanggan di Dealer Honda Arista antara lain *Revo, Supra, Beat, Vario, Sonic, PCX, CB,*

Scoopy, Megapro, Verza, dan CBR (dengan type yang beragam).

Sebagai dealer tentunya banyak informasi-informasi yang dibutuhkan oleh Honda Arista guna meningkatkan penjualan serta mengembangkan usahanya. Salah satu bentuk informasi penting yang dibutuhkan oleh dealer kendaraan bermotor adalah informasi tentang seberapa besar pola kesetiaan konsumen terhadap kendaraan bermotor yang dijual, serta bagaimana memprediksi merek kendaraan bermotor yang akan dibeli oleh konsumen. Loyalitas terhadap suatu merek (*Brand Loyalty*) merupakan ukuran keterikatan seorang konsumen pada suatu merek, dan informasi yang penting bagi dealer merupakan inti dari kekuatan suatu merek (*Brand Quality*). Penelitian ini dilakukan agar data penjualan kendaraan bermotor dapat diolah menjadi informasi yang penting bagi dealer penjual kendaraan bermotor. Bentuk informasi yang penting yang diteliti adalah pola kesetiaan konsumen terhadap merek kendaraan bermotor. Peneliti menerapkan metode *data mining* menggunakan Algoritma *Apriori*. Hasil penelitian dapat membantu perusahaan atau dealer kendaraan bermotor untuk mendapatkan informasi *brand loyalty* konsumen terhadap merek kendaraan yang

dijualnya dan membantu prediksi merek kendaraan berikutnya yang akan dibeli seorang konsumen.

METODOLOGI PENELITIAN

2.1. Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual[1].

Data mining telah menarik banyak perhatian dalam dunia sistem informasi dan dalam masyarakat secara keseluruhan dalam beberapa tahun terakhir, karena ketersediaan luas dalam jumlah besar data dan kebutuhan segera untuk mengubah data tersebut menjadi informasi yang berguna dan pengetahuan(Witten, 2012). Informasi dan pengetahuan yang diperoleh dapat digunakan untuk aplikasi mulai dari pasar analisis, deteksi penipuan, dan retensi pelanggan, untuk pengendalian produksi dan ilmu pengetahuan eksplorasi(Han&Kamber, 2007). Adanya ketersediaan data yang melimpah, kebutuhan akan informasi atau pengetahuan sebagai sarana pendukung dalam pengambilan keputusan baik bagi individu, organisasi, perusahaan dan pemerintahan[2]. Tugas dari data mining adalah untuk menghasilkan semua kaidah asosiasi pada suatu table transaksional, yang mempunyai nilai support lebih dari minimum support. Kaidah tersebut juga harus mempunyai support yang lebih besar dari confidence[3].

2.2. Algoritma Apriori

Apriori adalah algoritma kelas yang membantu mempelajari peraturan asosiasi. Ini bekerja melawan transaksi. Algoritma mencoba untuk menemukan himpunan bagian yang umum dalam kumpulan data. Ambang batas minimum harus dipenuhi agar asosiasi dapat dikonfirmasi[4].

Metode apriori dimana apriori termasuk jenis aturan asosiasi. Analisis asosiasi adalah teknik untuk menemukan aturan asosiatif antara suatu kombinasi item. Barang yang dianalisis yaitu barang yang dipesan bersamaan dengan barang lainnya atau dengan kata lain pemesanan lebih dari satu barang namun tetap melibatkan data pemesanan keseluruhan[5].

2.3. Tahapan Algoritma Apriori

Berikut merupakan tahapan perhitungan dengan *Data Mining* menggunakan Algoritma Apriori :

1. Pembentukan kandidat itemset, kandidat k-itemset dibentuk dari kombinasi (k-1)-itemset yang didapat dari iterasi sebelumnya. Satu cara dari algoritma apriori adalah adanya pemangkasan kandidat k-itemset yang subset-nya yang berisi k-1 item tidak

termasuk dalam pola frekuensi tinggi dengan panjang k-1.

2. Penghitungan support dari tiap kandidat k-itemset. Support dari tiap kandidat k-itemset didapat dengan menscan database untuk menghitung jumlah transaksi yang memuat semua item didalam kandidat k-itemset tersebut. Ini adalah juga ciri dari algoritma apriori dimana diperlukan penghitungan dengan cara seluruh database sebanyak k-itemset terpanjang.
3. Tetapkan pola frekuensi tinggi. Pola frekuensi tinggi yang memuat k item atau k-itemset ditetapkan dari kandidat k-itemset yang supportnya lebih besar dari minimum support.
4. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Bila tidak, maka k ditambah satu dan kembali bagian 1.

2.4. Support

Support adalah nilai penunjang atau presentase kombinasi sebuah item dalam database[6]. Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai item diperoleh dengan rumus berikut :

$$\text{Support}(A) = \frac{\text{Jumlah transaksi yang mengandung } A}{\text{Jumlah total transaksi}}$$

Gambar 1. Rumus Support

2.5. Confidence

Apriori adalah suatu algoritma untuk melakukan pencarian frequent itemset untuk mendapatkan association rules. Sesuai dengan namanya, algoritma ini menggunakan prior knowledge mengenai frequent itemset properties yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Apriori menggunakan pendekatan secara iterative yang disebut juga sebagai level-wise search di mana k-itemset digunakan untuk mencari (k+1)-itemset[7].

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai item diperoleh dengan rumus berikut :

$$\text{Support}(A, B) = P(A \cap B) = \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Jumlah total transaksi}}$$

Gambar 2. Rumus Support

2.6. Assosiate Rule

Analisis asosiasi atau *association rule* mining adalah teknik data mining untuk menemukan aturan asosiasi antara kombinasi item. Aturan asosiasi akan menggunakan data latihan, sesuai dengan pengertian data mining, untuk menghasilkan pengetahuan. Pengetahuan untuk

mengetahui item-item belanja yang sering dibeli secara bersamaan dalam suatu waktu. Aturan asosiasi yang berbentuk "if...then..." atau "jika...maka..." merupakan pengetahuan yang dihasilkan dari fungsi Aturan Asosiasi[8].

Setelah support dan Confidence sudah ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat untuk confidence dengan menghitung confidence asosiatif aturan asosiatif $A \Rightarrow B$. Nilai confidence dari aturan $A \Rightarrow B$ diperoleh dari rumus berikut :

$$Confidence(A \Rightarrow B) = \frac{\text{Jumlah transaksi yang mengandung A dan B}}{\text{Jumlah transaksi yang mengandung A}}$$

Gambar 3. Rumus Association Rule

HASIL DAN PEMBAHASAN

3.1. Analisis Data

Tahap ini melakukan analisa terhadap nilai-nilai objek yang akan diangkat, kemudian *value* atau nilai data tersebut akan diolah untuk mendapatkan informasi yang nantinya akan digunakan dalam tahap implementasi. Agar data yang akan diolah benar-benar relevan dengan yang dibutuhkan maka dilakukan Penggabungan data (*Integrasi Data*) kedalam suatu penyimpanan baru kemudian ditemukan atribut – atribut yang digunakan dalam data konsumen[9]. Adapun atribut-atribut di pakai adalah:

1. Pembelian dalam 1 tahun terakhir.
2. Item adalah Kategori untuk Merek kendaraan bermotor.

Tabel 1. Item Bought

NO	Item Bought
1	Revo, Supra, Beat, Vario, Sonic
2	PCX, CB, Scoopy, Megapro, Verza, CBR
3	Supra, Vario, Scoopy, Verza, Beat
4	Supra, Beat, PCX, Vario, Revo, Megapro, Scoopy
5	Scoopy, Verza, Beat, Megapro
6	PCX, Megapro, CB, CBR, Sonic, Supra, Vario, Beat
7	Revo, PCX, Sonic, CB, Scoopy, CBR, Megapro
8	Beat, Vario, Supra, Verza, PCX, Megapro
9	PCX, Supra, Beat, Scoopy, Sonic, CB, Vario, Verza
10	Sonic, PCX, Megapro, Supra, Scoopy, Verza, Beat, Vario, CBR
11	Supra, Sonic, PCX, Beat, Megapro, Revo, CB, Scoopy, Verza
12	PCX, Megapro, Verza, Supra, Scoopy, Beat, CBR, Vario

Ditentukan :

Minimum Support = 60 %

Minimum Confidence = 80 %

3.2. Menentukan Frequent Item Set (FIS)

Tabel 2. FIS

No	Item	Jumlah	Jumlah Seluruh Item	Hasil
1	Revo	4	12	33%
2	Supra	9	12	75%

3	Beat	9	12	75%
4	Vario	8	12	67%
5	Sonic	6	12	50%
6	PCX	9	12	75%
7	CB	5	12	42%
8	Scoopy	9	12	75%
9	Megapro	9	12	75%
10	Verza	8	12	67%
11	CBR	5	12	42%

Maka Item {Supra, Beat, Vario, PCX, Scoopy, Megapro, Verza} adalah FIS nya

Menentukan *Frequent Item Set (FIS)* besar dari minimum Support

1. Support(Supra \Rightarrow Beat) = Supra U Beat=8/12 = 67%
2. Support(Supra \Rightarrow Vario) = Supra U Vario=7/12= 58%
3. Support(Supra \Rightarrow PCX) = Supra U PCX=3/12 = 25%
4. Support(Supra \Rightarrow Scoopy) = Supra U Scoopy=5/12= 42 %
5. Support(Supra \Rightarrow Megapro) =Supra U Megapro= 2/12 = 17%
6. Support(Supra \Rightarrow Verza) = Supra U Verza=5/12=42%
7. Support(Beat \Rightarrow Vario) = Beat U Vario = 6/12 = 50%
8. Support(Beat \Rightarrow PCX) = Beat U PCX = 1/12 = 8%
9. Support(Beat \Rightarrow Scoopy) = Beat U Scoopy=2/12=12%
10. Support(Beat \Rightarrow Megapro) = Beat U Megapro = 3/12 = 25%
11. Support(Beat \Rightarrow Verza) = Beat U Verza = 4/12 = 33%
12. Support(Vario \Rightarrow PCX) = Vario U PCX = 1/12 = 8%
13. Support(Vario \Rightarrow Scoopy) = Vario U Scoopy = 1/12 = 8%
14. Support(Vario \Rightarrow Megapro) = Vario U Megapro = 1/12 = 8%
15. Support(Vario \Rightarrow Verza) = Vario U Verza = 3/12 = 25%
16. Support(PCX \Rightarrow Scoopy) = PCX U Scoopy = 7/12 = 58%
17. Support(PCX \Rightarrow Megapro) = PCX U Megapro = 8/12 = 67%
18. Support(PCX \Rightarrow Verza) = PCX U Verza = 5/12 = 42%
19. Support(Scoopy \Rightarrow Megapro) = Scoopy U Megapro = 3/12 = 25%
20. Support(Scoopy \Rightarrow Verza) = Scoopy U Verza = 6/12 = 50%
21. Support(Megapro \Rightarrow Verza) = Megapro U Verza = 4/12 = 33%

Berikut adalah data perhitungan:

Tabel. 3. Minimum Support

Supra=>Beat	8	12	67%
Supra=>Vario	7	12	58%
Supra=>PCX	3	12	25%
Supra=>Scoopy	5	12	42%
Supra=>Megapro	2	12	17%
Supra=>Verza	5	12	42%
Beat=>Vario	6	12	50%
Beat=>PCX	1	12	8%
Beat=>Scoopy	2	12	17%
Beat=>Megapro	3	12	25%
Beat=>Verza	4	12	33%
Vario=>PCX	1	12	8%
Vario=>Scoopy	1	12	8%
Vario=>Megapro	1	12	8%
Vario=>Verza	3	12	25%
PCX=>Scoopy	7	12	58%
PCX=>Megapro	8	12	67%
PCX=>Verza	5	12	42%
Scoopy=>Megapro	3	12	25%
Scoopy=>Verza	6	12	50%
Megapro=>Verza	4	12	33%

Dan hasil FIS besar dari Minimum Support adalah {(Supra, Beat), (PCX, Megapro)}.

3.3. Menentukan Minimum Confidence

- Confidence(Supra => Beat) = Support(Supra => Beat)/Support(Supra) = (8/12)/ (9/12) = 89%
- Confidence(Beat => Supra) = Support(Beat => Supra)/Support(Beat) = (8/12)/ (9/12) = 89%
- Confidence(PCX => Megapro) = Support(PCX => Megapro)/Support(PCX) = (8/12)/(9/12)=89%
- Confidence(Megapro => PCX) = Support(Megapro => PCX)/Support(Megapro) = (8/12)/ (9/12) = 89%

3.4. Menentukan Interesting Rule

Berikut adalah rumus untus Interesting Rule :

Body => Head [Support,Confidence]

Maka : Supra => Beat [67%, 89%]
 Beat => Supra [67%, 89%]
 PCX => Megapro [67%, 89%]
 Megapro => PCX [67%, 89%]

3.5. Menentukan Knowledge

Supra => Beat

- 67% dari semua transaksi, item Beat & Supra dibeli secara bersamaan
- Dari semua transaksi yang memberi otem Beat 89% membeli item Supra
- 67% dari semua transaksi, item PCX & Megapro dibeli secara bersamaan
- Dari semua transaksi yang memberi otem PCX 89% membeli item Megapro

KESIMPULAN

Berdasarkan hasil analisis dan pengujian yang telah dilakukan pada implementasi Data Mining menggunakan Algoritma Apriori, maka dapat diambil kesimpulan sebagai berikut :

- Hasil dari *association* yang berupa iformasi mengenai Honda merk apa saja yang dibeli secara bersamaan oleh konsumen, dapat digunakan sebagai bahan pertimbangan untuk menetapkan Strategi Pemasaran.
- Pada data transaksi penjualan yang ada pada Showroom Honda Arista Pematangsiantar, dapat dihasilkan tentang hubungan asosiasi antar Honda yang biasanya di beli oleh pelanggan secara bersamaan.

DAFTAR PUSTAKA

- P. Studi, S. Informasi, F. Ilmu, K. Universitas, and L. Kuning, "PENGUNAAN ALGORITMA APRIORI DATA MINING UNTUK MENGETAHUI TINGKATKESETIAAN KONSUMEN (BRAND LOYALTY) TERHADAP MEREK KENDERAAN BERMOTOR (STUDI KASUS DEALER HONDA RUMBAL)," vol. x, no. x, pp. 44–52, 2016.
- M. Badruk, "Algoritma asosiasi dengan algoritma apriori untuk analisa data penjualan," *Pilar Nusa Mandiri*, vol. XII, no. 2, pp. 121–129, 2016.
- D. Mining, "PENERAPAN ALGORITMA APRIORI DALAM POLA PEMINJAMAN BUKU (STUDI KASUS : PERPUSTAKAAN STIKES PRIMA JAMBI)," vol. 6, no. 01, 2017.
- P. Studi, S. Informasi, F. T. Informasi, and U. Stikubank, "Implementasi data mining menggunakan algoritma apriori 1,2," pp. 372–382, 2017.
- J. Ilmiah, I. Komputa, E. Volume, A. Issn, A. W. Kusumaatmaja, and J. D. Bandung, "Membangun Aplikasi E-Commerce Suku cadang Motor di Surya Mandiri Motor," *J. Ilm. Komput. dan Inform.*, vol. 1, 2013.
- F. Marisa and D. Purnomo, "Penerapan Algoritma Apriori Terhadap Data Penjualan di Toko Gudang BM," no. 35, pp. 1–5.
- A. Anas, "Untuk Mendapatkan Pola Peminjaman Buku Perpustakaan Smpn 3 Batanghari," *J. Ilm. Media SISFO*, vol. 10, no. 2, pp. 628–641, 2016.
- K. Tampubolon, H. Saragih, and B. Reza, "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan," *Maj. Ilm. Inf. dan Teknol. Ilm.*, vol. 1, no. 1, pp. 93–106, 2013, doi: 10.1017/S0079497X00014341.
- T. Limbong, "Peran dan Fungsi Komputer dalam Mendukung Prestasi Akademik Mahasiswa STMIK Budi Darma Medan," *Inf. dan Teknol. Ilm.*, vol. 3, no. 1, pp. 138–143, 2014.