

Implementasi Apriori Untuk Menentukan Pola Asosiasi Kerusakan Sparepart

Agustinus Agung Hario Wibisono ^{#1}, Ade Rahmat Iskandar ^{*2}, Atik Febriani ^{#3}

Institut Teknologi Telkom Purwokerto

Jl. D. I. Panjaitan No. 128, Purwokerto Selatan, Banyumas, Jawa Tengah, 53147

Purwokerto

¹ 15102044@st3telkom.ac.id

² ade.rahmat@ittelkom-pwt.ac.id

³ atik@ittelkom-pwt.ac.id

Accepted on April 11, 2019

Abstrak

MONDO vespa merupakan sebuah tempat service motor dan penjualan sparepart vespa. Tujuan dari penulisan ini adalah untuk menghasilkan suatu rancangan aplikasi yang dapat membantu untuk konsumen mengetahui sparepart apa saja yang harus diganti saat mempunyai masalah di motor vespa. Dalam merancang suatu sistem informasi untuk membantu konsumen dalam memilih sparepart yang akan diganti dibutuhkan suatu aplikasi yang gampang dan mudah diakses sehingga konsumen dapat terbantu dalam mengetahui sparepart. Metode penulisan yang digunakan adalah metode *Association Rules*, dan dibangun dengan bahasa pemrograman *Android Studio* dan untuk data menggunakan *SQLite*. Aplikasi ini dapat membantu konsumen dalam mengganti sparepart vespa dan mengetahui jenis sparepart.

Kata Kunci: Data Mining, Market Basket Analysis, *Association Rules*, Algoritma Apriori

I. PENDAHULUAN

Zaman sekarang mulai banyak pengguna vespa di kota Purwokerto, dimana anak muda sudah mulai berkembang menghidupkan mesin 2 tak tersebut, tetapi mereka hanya tau dan menggunakan saja tidak tau jika ada trouble di jalan atau di tempat lain. Disini penulis ingin membuat aplikasi yang memudahkan pengguna motor vespa yang belum tau tentang kerusakan dan *sparepart* apa saja jika mereka mengalami kerusakan. Aplikasi ini sangat gampang digunakan dan dapat dipakai oleh siapa aja, aplikasi ini juga hanya di download dan tidak perlu update lagi, aplikasi ini dapat di download di play store punya android. Aplikasi ini punya sistem untuk memberikan jawaban saat pengguna vespa sudah tidak merasa enak dengan motor mereka.

II. METODE PENELITIAN

1.1 Data Mining

Data *Mining* adalah penemuan informasi baru dengan mencari aturan tertentu dari sejumlah data yang sangat banyak. Data *Mining*, Sering disebut juga sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi

pengumpulan, pemakaian data, historis untuk menemukan ketarutan hubungan itemset data yang sangat banyak. Karakteristik Data *Mining* sebagai berikut:[6]

1. Data Mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang belum diketahui .
2. Data Mining biasa menggunakan data yang sangat banyak. Biasanya data yang sangat banyak digunakan untuk membuat hasil lebih dipercaya.
3. Data mining berguna untuk membuat keputusan yang benar,terutama dalam strategi.

Menurut Turban, dkk menegaskan “Data Mining merupakan proses yang menggunakan teknik statistik, matematik, kecerdasan buatan, dan *machine learning* untuk menemukan dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai data yang banyak” [7].

Sedangkan menurut Budi Santosa menegaskan bahwa “Keluaran dari data mining bisa dipakai untuk memperbaiki pengambilan keputusan di tahun berikutnya” [7].

Adapun fungsi dari data mining itu sendiri adalah untuk:[8]

1. Deskripsi. Pola data sering dideskripsikan. Deskripsi tersebut sangat membantu dalam menjalankan pola yang terjadi. Model *data mining* harus transparan ,akan tetapi pola datanya harus jelas.
2. Estimasi. Estimasi mirip dengan klasifikasi kecuali variable target-nya numerik ketimbang kategori. Model yang dibangun menggunakan *record* yang lengkap,yang menyediakan variable target dan predictor. Observasi yang baru, nilai ditentukan oleh estimasi nilai variable target yang ditentukan dan berdasarkan nilai predictor.
3. Prediksi. Prediksi sama dengan klasifikasi dan estimasi, kecuali dalam prediksi,hasil terjadi di masa yang akan datang.
4. Klasifikasi. Dalam Klasifikasi, Variable merupakan target yang sudah ditentukan. Model *data mining* memeriksa set *record* yang banyak, tiap *record* mempunyai informasi variable target dan set input atau variable predictor.
5. Clustering. Clustering merupakan pengelompokkan *record*, observasi, atau kasus ke dalam kelas-kelas objek yang sangat mirip. Clustering berbeda dengan klasifikasi dimana dalam clustering tidak mempunyai target variable. Clustering mencoba menyegmentasi seluruh set data ke dalam *subgroup* atau *cluster* yang relatif homogen,dimana mempunyai kesamaan antar *record* dalam *cluster* dimaksimasi dan kemiripan *record* di luar *cluster* diminimasi.
6. Asosiasi. Asosiasi merupakan suatu tugas untuk menemukan atribut-atribut yang terjadi jika output bersamaan. Tugas asosiasi mencoba untuk menemukan suatu aturan untuk mengkuantifikasi hubungan antara dua atau lebih atribut. Aturan asosiasi mempunyai bentuk “*IF antecedent, then consequent*”, bersama dengan ukuran *support* dan *confidence* yang berhubungan dengan aturan.

1.2 Market Basket Analysis

Menurut Jiawei Han dan Micheline Kamber menegaskan “*Market basket analysis* adalah suatu metodologi untuk melakukan analisis *buying* habit konsumen dengan menemukan asosiasi antar jumlah item yang berbeda,yang diletakkan konsumen dalam keranjang belanja yang di beli pada suatu barang tertentu. Tujuan *market basket analysis* adalah untuk mengetahui suatu produk yang mana akan dibeli secara bersamaan” [9].

1.3 Association Rule

Menurut Kusriani dan Emha taufiq Luthfi menegaskan “*Association rule mining* adalah teknik data *mining* untuk menemukan aturan asosiasi antara gabungan item”[9]. *Interestingness measure* yang dapat dipakai dalam data mining Yaitu :

- a. *Support*, adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan barang.

b. *Confidence*, adalah suatu ukuran yang menunjukkan hubungan antar dua atau lebih item secara *conditional* (Berdasarkan Kondisi tertentu).

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap, yakni melakukan proses analisis pola frekuensi tinggi (*Frequent pattern*) dan Proses pembentukan aturan asosiasi [9].

Adapun aturan asosiasi ini adalah *Affinity analysis* melakukan studi atribut atau karakteristik yang berjalan secara bersamaan. Metode *affinity analysis* atau dikenal sebagai *market basket analysis*, mencoba menemukan asosiasi antar atribut, mendapatkan aturan yang mengkuantifikasi hubungan antara dua atribut atau banyak atribut, Aturan asosiasi berbentuk “*IF antecedent, then consequent*”, dilengkapi oleh support dan confidence sesuai aturan tersebut [8].

Ada dua metode utama dalam merepresentasikan tipe data dalam *market basket analysis* : format data barang dan format data tabular. Format data barang membutuhkan dua *field*, *field ID* dan *field isi*, dimana tiap *record* merepresentasikan hanya satu item. Tabel 2.2 merupakan tabel format barang.

Tabel 2.2. Format Data Transaksi

| Field ID Barang | Field Isi |
|-----------------|-----------|
| 1 | Item 1 |
| 1 | Item 1 |
| 2 | Item 1 |
| - | - |
| - | - |
| 150 | Item 1 |

Dalam format data tabular ,tiap *record* mempresentasikan transaksi terpisah , dimana *flag field* 0/1 sebanyak jumlah itemnya.

Tabel 2.3. Format Data Tabular

| Barang | Item 1 | Item 2 | ... | Item n |
|--------|--------|--------|-----|--------|
| 1 | 1 | 1 | 0 | 0 |
| 2 | 0 | 0 | 0 | 1 |
| - | - | - | - | - |
| - | - | - | - | - |
| 10 | 1 | 1 | 0 | 1 |

Penghitungan dari *Support*, *Confidence*, *Frequent Itemset* : misalkan D adalah set item Barang, dimana tiap Barang T di D merepresentasikan set item yang ada di I. Aturan asosiasi berbentuk if A, then B ($A \Rightarrow B$), dimana *antecedent* A dan *consequent* B subset dari I, dan A dan B *mutually exclusive*.

Support untuk aturan asosiasi tertentu $A \Rightarrow B$ adalah proporsi Barang di D yang berisi baik A dan B.

$$Support = P(A \cap B) = \frac{\text{number of transactions containing both A and B}}{\text{total number of transactions}} \dots\dots(1)$$

Confidence c aturan asosiasi tertentu $A \Rightarrow B$ adalah ukuran akurasi aturan, yang dapat ditentukan dengan presentasi barang di D yang berisi A yang juga berisi B.

$$Confidence = P(B | A) = \frac{P(A \cap B)}{P(A)} = \frac{\text{number of transactions containing both A and B}}{\text{number of transactions containing A}} \dots\dots(2)$$

Aturan yang diinginkan adalah aturan yang mempunyai *support* yang lebih tinggi atau *confidence* yang lebih tinggi, dan biasanya keduanya. Aturan yang kuat adalah aturan yang memenuhi atau melebihi kriteria *support* dan *confidence* minimum.

Itemset adalah set item yang berisi di dalam I, dan k-itemset adalah item yang berisi k items. *Itemset frequency* adalah jumlah barang yang berisi itemset tertentu. *Frequent itemset* adalah itemset yang terjadi paling sedikit pada jumlah tertentu, mempunyai $frequency \geq (\text{phii})$. Misalkan, $= 4$, itemset yang terjadi lebih dari 4 atau disebut *frequency* [8].

1.4 Algoritma Apriori

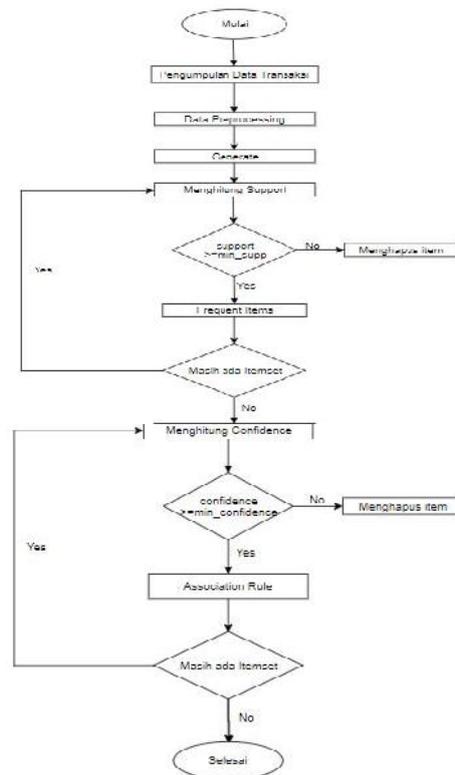
Apriori adalah algoritma untuk melakukan pencarian *frequent itemset* untuk mendapatkan *association rules*. Algoritma ini menggunakan *prior knowledge* mengenai *frequent itemset properties* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Apriori menggunakan pendekatan secara iterative yang disebut sebagai *level-wise search* dimana k-itemset digunakan untuk mencari (k+1)-itemset [5].

Pola *Frequent* tinggi adalah pola-pola item dalam database yang memiliki frekuensi atau *support* di atas ambang batas yang telah di tentukan disebut dengan minimum *support*. pola dari frekuensi tinggi ini digunakan untuk menata aturan asosiatif dan juga beberapa teknik data mining lainnya [5].

Prinsip algoritma apriori adalah :

1. Kumpulkan jumlah item tunggal, dapatkan item yang tujuan sama.
2. Dapatkan *candidate pairs*, hitung \rightarrow *large pairs* dari item-item
3. Dapatkan *candidate triplets*, hitung \rightarrow *large triplets* dari item-item dan seterusnya.
4. Sebagai petunjuk : setiap subset dari sebuah *frequent itemset* harus menjadi *frequent*.

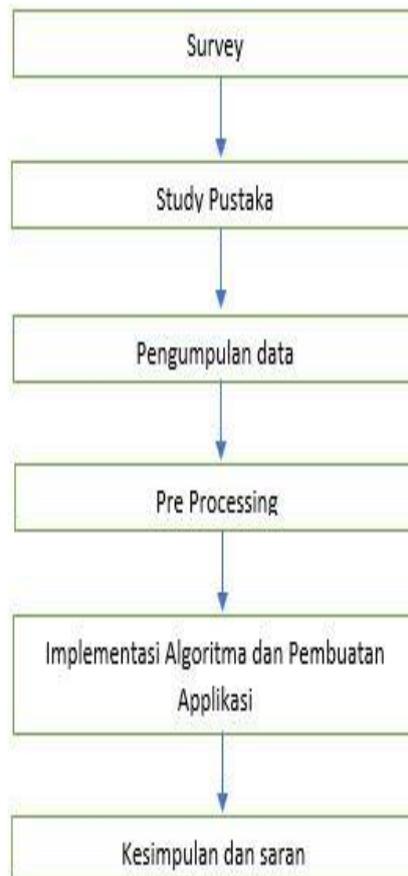
Dua proses utama dalam algoritma apriori merupakan langkah yang harus dilakukan untuk mendapatkan itemset. Algoritma Apriori mudah dipahami dan kekuarangnya apriori harus melakukan *scanning database* berulang kali untuk setiap kombinasi item. Banyaknya waktu yang dibutuhkan untuk melakukan scanning database selain itu dibutuhkan *generate candidate* yang besar untuk mendapatkan kombinasi item dari *database* [5].



Gambar 2.5 Flowchart algoritma Apriori

III. HASIL DAN PEMBAHASAN

Langkah-langkah dalam penyelesaian kerja penelitian ,masalah yang dibahas akan dibuat flowchart seperti dibawah ini :



Gambar 3.1 Tahapan – tahapan penelitian

3.1.1 Survey

Survey yang dilakukan dengan cara mengisi kuesioner kepada para pengguna vespa di Purwokerto, ada beberapa pertanyaan yang sudah dibuat. Pertanyaan ini dibuat untuk orang-orang yang mengendarai sepeda motor vespa saja. kuesioner ini dilakukan di bengkel Mondo Vespa Autocare agar mendapatkan informasi yang jelas seberapa banyak mereka mengetahui tentang *sparepart* vespa.

3.1.2. Studi Pustaka

Setelah melakukan survei dan menemukan permasalahan, kemudian dilakukan studi pustakan dengan membaca makalah yang membahas metode untuk pencarian pola asosiasi.

3.1.3. Pengumpulan Data

Pengumpulan data dilakukan untuk mengetahui jumlah sparepart yang ada dan data transaksi, hal ini dilakukan untuk mengambil data apa saja yang sering diganti saat sedang *service* vespa di bengkel Mondo Vespa Autocare, dengan cara wawancara langsung dan meminta data ke bengkel tersebut. Setelah data sudah diperoleh dalam bentuk manual data, kemudian ke microsoft excel untuk dibuat tabel dan dihitung dengan cara mencari nilai *support* dan *confidence* seperti yang sudah ditulis di bab sebelumnya.

3.1.4. Pre Processing

Pada tahap ini penulis mengolah data yang sudah didapatkan agar bisa diimplementasikan pada algoritma dan menerapkannya pada pembuatan aplikasi.

3.1.5. Implementasi Algoritma dan Pembuatan Aplikasi

Setelah data melalui tahap Pre Processing atau tahap mengolah data, maka dilanjutkan pada pengimplementasian menggunakan algoritma apriori, dan disertai dengan pembuatan aplikasi, yang selanjutnya akan dilakukan analisis dari hasil implementasi algoritma dan penerapannya pada aplikasi.

3.1.6. Analisis Hasil

Setelah diimplementasikan pada algoritma dan telah dibuatkan aplikasi, maka penulis menganalisa hasil, apabila hasil sudah sesuai dengan yang diharapkan maka dilanjutkan pada tahap kesimpulan dan saran.

3.1.7. Kesimpulan dan Saran

Pada tahapan ini penulis menyimpulkan dari hasil data yang sudah diolah dan diimplementasikan, agar bisa menjadi acuan kedepannya bagi penelitian selanjutnya.

IV. KESIMPULAN DAN SARAN

Kesimpulan

Dalam penelitian dan pengimplementasikan data mining pada pengambilan keputusan dalam menganalisis part vspa menggunakan algoritma apriori, maka didapatkan beberapa kesimpulan:

1. Algoritma apriori telah diterapkan untuk membuat aplikasi vespa antar part pada transaksi yang terjadi di meicrosoft excel. Sehingga hasil yang diperoleh dapat menjadi acuan untuk mengambil keputusan dalam menganalisis aplikasi yang dibuat dengan android studio guna membantu dalam kerusakan pada vespa.
2. Implementasi dilakukan dengan menggunakan microsoft excel sebagai penyimpanan data hasil kombinasi sementara. Selain itu penghitungan yang menggunakan supp dan confidence juga dipakai untuk mengkombinasikan antar part dalam itemset.

Saran

Dari penelitian yang dilakukan, dapat disarankan :

1. Dalam pembuatan aplikasi ini masih terdapat beberapa kekurangan khususnya dari faktor desai,. Sehingga dibutuhkan masukan yang sangat mendukung bagi pengembangan dan kemajuan aplikasi ini.
2. Dalam mengimplementasikan algoritma Apriori untuk menemukan aturan asosiasi part dalam aplikasi seperti yang dibangun diperlukan data yang banyak dari pelanggan vespa.

REFERENSI

- [1] Frismadani Anggita Priyana, "Data Mining Asosiasi Untuk Menentukan Cross-Selling Produk Menggunakan Algoritma Frequent Pattern-Growth Pada Koperasi Karyawan PT. Phapros Semarang," vol. 1, no. 2, pp. 0–12, 2015.
- [2] B. N. Hapsari Dita Anggraini, Ragil Saputra, "Aplikasi Data Mining Analisis Data Transaksi Penjualan Obat Menggunakan Algoritma Apriori (Study Kasus di Apotik Setya Sehat Semarang)," vol. 4, pp. 1–8, 2012.
- [3] G. J. Tamaela, "Evaluasi Kinerja Algoritma," no. Mueller 1995, pp. 38–45, 2007.
- [4] D. Haryanto, Y. Oslan, and D. Dwiyana, "Implementasi Analisis Keranjang Belanja Dengan Aturan Asosiasi Menggunakan Algoritma Apriori Pada Penjualan Suku Cadang Motor," pp. 81–94.
- [5] A. Azwar, "Analisa Algorithma Apriori Untuk Mendapatkan Pola Peminjaman Buku Perpustakaan Smpn 3 Batanghari," *J. Ilm. Media SISFO*, vol. 10, no. 2, pp. 1978–8126, 2016.
- [6] Paul Beynon-Davies, "Database Systems Third Edition," vol. 1, no. 1, pp. 1–12, 2004.
- [7] K. Tampubolon, H. Saragih, B. Reza, K. Epicentrum, A. Asosiasi, and A. Apriori, "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan," *Inf. dan Teknol. Ilm.*, pp. 93–106, 2013.
- [8] Y. W. Yusuf, P. F. Rian, and T. Gerry, "Penerapan Data Mining Dalam Penentuan Aturan Asosiasi Antar Jenis Item," *Semin. Nas. Apl. Teknol. Inf. 2006 (SNATI 2006)*, vol. 2006, no. Snati, pp. E53–E56, 2006.
- [9] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth);," *Telematika*, vol. 4, no. 1, pp. 118–132, 2012.
- [10] C. Chandra, "Perancangan Program Aplikasi Market Basket Analysis untuk mendukung persediaan Barang Dengan Metode Fuzzy C-Covering," *Math. Stat. Dep. Sch. Comput. Sci. Binus Univ.*, pp. 241–253.